# (19) World Intellectual Property Organization

International Bureau





(43) International Publication Date 1 July 2004 (01.07.2004)

**PCT** 

## (10) International Publication Number WO 2004/055216 A2

(51) International Patent Classification7:

C12Q 1/70

(21) International Application Number:

PCT/US2003/039722

(22) International Filing Date:

12 December 2003 (12.12.2003)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data: 60/433,303

13 December 2002 (13.12.2002) US

(71) Applicant (for all designated States except US): FOX CHASE CANCER CENTER [US/US]; 333 Cottman Avenue, Philadelphia, PA 19111 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ZHU, Qing [CN/US]; 1504 Liberty Court, North Wales, PA 19454 (US). GUO, Ju-Tao [CN/US]; 26 Township Line Road, Apt. A7, Elkins Park, PA 19027 (US). SEEGER, Christoph [US/US]; 407 Waring Road, Elkins Park, PA 19027 (US).

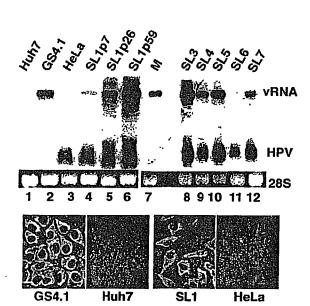
- (74) Agents: RIGAUT, Kathleen, D. et al.; Dann, Dorfman, Herrell & Skillman, Suite 2400, 1601 Market Street, Philadelphia, PA 19103-2307 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: REPLICATION OF HEPATITIS C VIRUS IN NON-HEPATIC EPITHELIAL AND MOUSE HEPATIC CELLS



(57) Abstract: Cells and cell lines which replicate HCV of non-hepatic human and non human origin are disclosed. Also provided are methods of using such cells and cell lines to identify anti-HCV agents for the treatment of HCV infection.

WO 2004/055216 A2 ||||||||||

# REPLICATION OF HEPATITIS C VIRUS IN NON-HEPATIC EPITHELIAL AND MOUSE HEPATIC CELLS

Qing Zhu

Ju-Tao Guo

Christoph Seeger

This application claims priority to U.S. provisional application 60/433,303, filed December 13, 2002, the entire contents of which are incorporated by reference herein.

### GOVERNMENT RIGHT

5

10

15

Pursuant to 35 U.S.C. Section 202(c), it is acknowledged that the United States government has certain rights in the invention described herein, which was made in part with funds from the National Institutes of Health Grant No. AI48046.

#### FIELD OF THE INVENTION

20 This invention relates to the fields of molecular biology and pathology. Novel animal cell lines and non-hepatic human epithelial cell lines for the replication of hepatitis C virus (HCV), as well as methods for screening for anti-HCV drugs or HCV receptors using these cell lines are disclosed. Furthermore, adaptive sequence mutations in the HCV genome, which permit replication in non-human, and non-hepatic cell lines are also provided.

### BACKGROUND OF THE INVENTION

- 30 Several publications and patent documents are cited in this application in order to more fully describe the state of the art to which this invention pertains. The disclosure of each of these citations is incorporated by reference herein.
- 35 Hepatitis C virus (HCV) is an enveloped, positive

stranded RNA virus that belongs to the Flaviviridae, a family of viruses including human pathogens such as yellow fever virus, dengue virus and West Nile virus (Q. L. Choo et al., Science 244, 359-62 (1989)). Although 5 broad tissue and species tropisms are hallmarks of these viruses, HCV replication has so far only been detected in human and chimpanzee livers. Moreover, for reasons that are not yet understood, HCV RNA levels in infected liver tissue are extremely low, generally below one copy of RNA per cell and hence, can only be detected with PCR, making it difficult to determine whether secondary sites for viral replication exist (J. Boisvert et al., J Infect Dis 184, 827-35 (Oct 1, 2001); R. E. Lanford, et al., J Virol 69, 8079-83 (1995)).

HCV encodes a polyprotein that is processed 15 proteolytically into ten polypeptides (K. E. Reed, C. M. Rice, Curr Top Microbiol Immunol 242, 55-84 (2000)). Three of them are structural proteins required for capsid formation (core) and assembly into enveloped viral particles (E1 and E2). Four of them are enzymes 20 including cysteine and serine proteases (NS2 and NS3), an ATP dependent helicase (NS3) and a RNA-directed RNA polymerase (NS5B). The functions of the remaining three polypeptides, p7, NS4B, and NS5A, for viral replication are not yet known. For study of replication of HCV in 25 tissue culture cells, the structural proteins can be replaced with a selectable marker, such as the neomycin phosphotransferase. See for example Figure 2, left panel of Lohman et al. (V. Lohmann et al., Science 285, 110-3 (1999)). Replication of such subgenomic HCV replicons in 30 tissue culture cells has so far only been demonstrated in the human hepatoma cell line Huh7, consistent with the narrow host and tissue tropism of HCV infections.

HCV infection poses a significant public health problem. Approximately 3% of the world's population has

35

persistent HCV infection. In 1989, the virus was identified as the major aetiological agent responsible for post-transfusion non-A and non-B hepatitis.

Following primary HCV infection, persistent viraemia and chronic hepatitis develop in the majority of cases.

Efforts to elucidate the mechanisms behind viral persistence and hepatocellular damage have been frustrated by the lack of a reliable cell culture system for viral propagation in vitro. In addition, as the chimpazee is the only experimental animal susceptible to HCV infection, progress in research is hampered by the lack of a small animal model to facilitate pathophysiological studies as well as the evaluation of antiviral treatment and vaccine strategies.

Furthermore, although the initial HCV infection is 15 asymptomatic, subsequent clinical manifestations of HCV induced liver disease include fibrosis, cirrhosis, and hepatocellular carcinoma (Alter, H. J., and L. B. Seeff. 2000. Semin. Liver Dis. 20:17-35). Combination antiviral therapy with alpha interferon (IFN- $\alpha$ ) and ribavirin, a 20 purine nucleoside analogue, arrests disease progression and can lead to sustained recovery in only 45 to 80% of treated patients (Di Bisceglie, A. M., and J. H. Hoofnagle. 2002. Hepatology 36:S121-S127). Additionally, response to IFN- $\alpha$  therapy can vary significantly 25 depending on the viral genotype, ranging from 30 to 40% for genotype 1 to as high as 80% for genotypes 2 and 3. This suggests that viral determinants also play an important role in regulating the cellular IFN response against HCV (Kinzie, J. L., et al., 2001. J. Viral 30 Hepatitis 8:264-269; McHutchison, J. G., et al., 1998. N. Engl. J. Med. 339:1485-1492). The parameters determining the success or failure of antiviral therapy are not understood, and their identification represents a major 35 challenge in HCV biology.

Therefore, there is a desperate need for non-hepatic cell culture systems, and small animal models for the identification and characterization of anti-viral agents for the prevention and treatment of HCV infection.

Additionally, there is a need in the art to elucidate the

Additionally, there is a need in the art to elucidate the mechanism of HCV inhibition by IFN- $\alpha$ , so that other treatments may be found.

#### SUMMARY OF THE INVENTION

20

35

10 The present invention provides HCV replicating cells and cell lines derived from human non-hepatic cells or non-human cells. According to one embodiment of the invention, the cells are human epithelial cells of non-liver origin, such as, HeLa cells. According to another embodiment of the invention, the cells capable of replicating HCV are hepatoma and hepatocyte cells of mouse origin, such as, Hepal-6 cells, or AML12 cells respectively.

The present invention also provides a non-human host animal comprising cells infected with HCV. In one embodiment of the invention, the host animal is a mouse. In another embodiment of the invention, the cells infected with HCV are mouse hepatoma cells.

Also provided by the present invention are methods

for producing human non-hepatic cells or non-human cells
that are capable of replicating HCV, and cell lines
comprising the same. Such methods include transfection
with total HCV RNA or an HCV replicon which comprises one
or more adaptive mutations which facilitate replication

in a cell of interest.

The present invention further provides methods for screening an agent that modulates HCV replication by incubating the agent with the aforementioned cells or administering the agent to the aforementioned host animal comprising cells replicating HCV and assessing said agent

for modulation of HCV replication. Such agents may inhibit or enhance production of HCV. These agents may be cytopathic or non-cytopathic to HCV infected cells. Agents which activate aspects of the JAK/STAT pathway may also be screened using the cells and cell lines of the invention.

Also provided by the present invention are HCV derived polynucleotides comprising adaptive mutations. The present inventor has discovered that these mutations are associated with expanded tropism of HCV.

Additionally, the present invention provides polypeptides encoded by the mutated HCV polynucleotides described above.

#### 15 BRIEF DESCRIPTION OF THE DRAWINGS

10

Figures 1A and 1B are a Northern blot (Fig. 1A) and micrographs showing replication of HCV subgenomic replicons in HeLa cells. (Fig. 1A) Detection of HCV viral RNA. Total RNA (5 µg) was isolated from HeLa cell lines (SL1, SL3-7) that were established from G418 20 resistant cell colonies and analyzed by Northern blot analysis. Blots were hybridized with radiolabeled RNA probes corresponding to the HCV NS5 region to detect viral RNA (vRNA) and the ΔE1 region of human papilloma virus (HPV) present in HeLa cells. In vitro transcribed 25 HCV RNA (1 ng, + 5µg total RNA from Huh7 cells, lane 7) served as a marker (M) and control for the hybridization reaction and 28S ribosomal RNA as a control for the amount of RNA present in each sample analyzed. GS4.1 is a 30 Huh7 derived cell line expressing HCV subgenomic replicons. RNA in SL1 cells was analyzed from cells harvested at the indicated passage (p). SL3-7 were analyzed at p3 (Fig. 1B). Immunohistochemical analysis of HCV replication in HeLa cells. Expression of NS5A in GS4.1 and SL1 cells (p26) was detected with a monoclonal 35

antibody bound to fluorescent isothiocyanate (FITC) - conjugated antibody. Parental Huh7 and HeLa cells served as controls.

Figure 2 diagrams sequence analysis of HCV subgenomes in HeLa and mouse hepatoma Hepal-6 cells. 5 (Left section) The physical map of HCV subgenomic RNA include the positions of the first amino acid NS3 and the last residue of the polyprotein. The internal ribosomal entry site for the translation of the NS genes is indicated (EMCV-IRES). (Right section) Mutations causing 10 amino acid changes identified in cDNAs isolated from subgenomic replicons present in GS4.1 and the indicated HeLa (SL1 (p26); SL2, p5) and Hepa1-6 cell lines (MH1 (p12), MH2 (p4), MH4 (p4)) are depicted with horizontal bars. Four independent clones were sequenced from each 15 PCR fragment that was amplified from cDNAs obtained from total RNA purified from the indicated cell lines. Mutations present in more than one cell line are indicated by amino acid position. Mutations that occurred in 50% of the clones analyzed are marked with an 20 asterisk. Mutations that occurred in only one of four clones analyzed were not included in the figure. A deletion identified in cDNA clones obtained from SL1 cells spanning amino acids 2371 to 2413 is indicated ( $\Delta$ ).

Figures 3A and 3B are a Northern blot and micrographs showing replication of HCV subgenomes in mouse hepatoma cells. (A) Detection of HCV RNA. RNA in Hepa1-6 cell lines (MH1-5) that were established from G418 resistant cell colonies and analyzed as described in the legend to Figure 1 except that a radiolabeled probe specific to mouse albumin cDNA was used in lieu of the probe against HPV. MH4-5 were analyzed at p4 (B). Immunohistochemical analysis of HCV replication in MH1 cells (p3). Expression of NS5A was detected as described in Figure 1. Hepa1-6 cells served as a

25

30

35

negative control.

5

10

15

30

35

Figure 4 is a Northern blot showing detection of HCV RNA in mouse hepatocyte cells. AML12 cells were transfected with total RNA from GS4.1 cells, which express subgenomic HCV replicons, and HFL cells, which express full length HCV genomes. 5 micrograms of total RNA was isolated from AML12 cell lines (MA6-5 to MA6-8, and MAC1) that were established from G418 resistant cell colonies and analyzed by Northern blot analysis. sgRNA indicates subgenomic RNA and flRNA full length genomic RNA.

Figures 5A and 5B are two Northern blots showing the antiviral activity of the HCV RNA polymerase inhibitor 2'-C-methyladenosine (2CMA). GS4.1 (Huh7) cells (5A) and SL1 (HeLa) cells (5B) were treated with 10 µM 2CMA. The cells were harvested at the indicated time points. Total cellular RNA was extracted and viral RNA (vRNA) analyzed by Northern blot analysis.

Figures 6A and 6B are two graphs depicting antiviral activity of the HCV RNA polymerase inhibitor 5-OH-cytidine. GS4.1 (Huh7) cells and SL1 (HeLa) cells were treated with the indicated amounts of 5-OH-cytidine. The DNA polymerase inhibitor 5-OH-deoxy-cytidine was used as a negative control. The cells were harvested 72 hours after incubation with the drugs. Total cellular RNA was extracted and viral RNA analyzed by Northern blot analysis. The intensity of the bands corresponding to HCV RNA was determined with a Fuji phosphoimager.

Figures 7A and 7B are a Northern blot and a graph showing antiviral activity of IFN- $\alpha$  in Huh 7 (GS4.1) and HeLa (SL1) cell lines containing HCV replicons. (7A) Viral RNA (vRNA) levels present in GS4.1 and SL1 cells incubated with 0, 0.1, 0.3, 1, 3, 10, 30, and 100 IU of IFN- $\alpha$ /ml (lanes 1 to 9 and 12 to 18) and with 0.01 and 0.03 IU/ml (lanes 10 and 11) for 72 h were determined by

Northern blot analysis. rRNA (28S rRNA) served as a control for the amount of RNA loaded per lane. (7B) Amounts of HCV RNA were determined with a Fuji phosphorimager, and the values were plotted as the percentages of the values obtained with untreated cells in lanes 1 and 9.

Figures 8A and 8B are micrographs and histograms showing that IFN-  $\alpha$  induces apoptosis of SL1 cells. (8A) Annexin V-FITC staining. SL1 cells grown on glass coverslips were left untreated (upper left) or treated with 100 IU of IFN-  $\alpha$  /ml for 6 h (upper right) or 20 h (lower left) or with 100 IU of IFN-  $\alpha$  /ml and 20  $\mu M$ caspase inhibitor ZVAD-FMK for 20 h (lower right). Cells were then processed for annexin V FITC staining and viewed with a fluorescence microscope. (8B) Flow cytometry analysis. SL1 cells were left untreated (upper left) or treated with 100 IU of IFN-  $\alpha$  /ml (upper right) or with IFN-  $\alpha$  and 20  $\mu M$  ZVAD-FMK (lower left) for 24 h. To inhibit viral replication, SL1 cells were incubated at 39°C for 60 h and then treated with 100 IU of IFN-  $\alpha$  /ml for 24 h (lower left). Cells were harvested and processed for annexin V-FITC staining and analyzed by flow cytometry. The percentages of FITC-positive cells are indicated.

10

15

20

Figures 9A and 9B are Northern blots and a graph showing a comparison of IFN-  $\alpha$  responses against HCV and flavivirus Kunjin virus replicons in HeLa cells. (9A) SL1 and KUNCD20 cells were incubated with 0, 0.01, 0.04, 0.16, 0.625, 2.5, 10, 40, and 160 IU of IFN-  $\alpha$  /ml (lanes 1 to 9 and 10 to 18, respectively) for 72 h, and viral RNA levels were determined by Northern blot analysis with a plus-strand-specific RNA probe for the neomycin phosphotransferase II gene. Mx-1 mRNA served as a control for IFN-  $\alpha$  -induced gene expression.  $\beta$ -Actin mRNA served as a control for the amount of RNA loaded per lane. (9B)

The amounts of HCV and Kunjin virus replicon RNA (arbitrary units) were determined with a Fuji phosphorimager.

Figure 10 shows an overview of the interactions between a virus and the IFN system. Replication of 5 viruses in cells produces dsRNA and viral proteins, which activate PKR and OAS/RNase L antiviral pathways and also signal to the promoter of the IFN- $\beta$  gene by activating transcription factors IRF3, NF-kB, and ATF2. Secreted IFN binds to its receptor and activates receptor-associated 10 Jak kinases, leading to the formation of the trimeric transcription factor ISGF3, which binds to the IFNstimulated response element (ISRE) on promoters of IFNstimulated genes. Among the products of the several hundred genes induced by IFN, PKR, OAS/RNase L, and Mx 15 are the best-characterized antiviral proteins, which inhibit different stages of viral replication and induce apoptosis of virally infected cells.

Figures 11A-11D are Northern blots and two graphs showing inhibition of the IFN-  $\alpha$  response by genistein 20 and the V protein of HPIV2. (11A) GS4.1 cells were incubated with 100 µg of genistein/ml for 2 h and then with 100 IU/ml IFN- $\alpha$  for an additional 24 h. Viral RNA levels were determined by Northern blot analysis. The cells were harvested at the indicated time points, and 25 Mx-A mRNA and viral RNA levels were determined by Northern blot analysis. Ribosomal 28S RNA was used as a control for the amount of RNA loaded on each lane. (11B) The amounts of viral RNA were measured with a phosphorimager and plotted as percentages of the values 30 obtained with untreated cells. (11C) GS4.1 cells were transfected with pCMV-E3L and pEF-HA-HPIV2 and treated with IFN- $\alpha$  at the indicated concentrations for 3 days. HCV RNA was subjected to Northern blot analysis. (11D) Viral RNA levels were determined with a Fuji 35

phosphorimager and plotted as the percentages of the values obtained with untreated cells.

Figures 12A and 12B show the dsRNA response in parental Huh7 and HeLa cells and HCV replicon-containing GS4.1 and SL1 cells. (12A) Phosphorylation of eIF-2 $\alpha$ . 5 Huh7, GS4.1, HeLa, and SL1 cells were left untreated (lanes 1, 5, 9, and 13) or treated with 100 IU of IFN- $\alpha$ /ml for 12 h (lanes 2, 4, 6, 8, 10, 12, 14, and 16) and then transfected with poly(I:C) and incubated for 3 h (lanes 3, 4, 7, 8, 11, 12, 15, and 16). eIF-2 $\alpha$  -P and 10 total eIF-2  $\alpha$  were determined by Western blots analysis with a monoclonal antibody specific for eIF-2α -P and an antibody specific for total eIF- $2\alpha$  protein. (12B) Induction of IFN- $\beta$  mRNA by dsRNA and IFN- $\alpha.$  Parental Huh7 and HeLa cells and HCV replicon-containing GS4.1 and SL1 15 cells were left untreated (lanes 1, 5, 9, and 13) or treated with 100 IU of IFN- $\alpha$  /ml (lanes 3, 4, 7, 8, 11, 12, 15, and 16) for 12 h and then transfected with poly(I:C) (lanes 2, 4, 6, 8, 10, 12, 14, and 16) for 3 h. An RNase protection assay was performed with probes 20 specific for IFN- $\beta$  and  $\beta$ -actin mRNAs.

Figures 13A-13C show dose-dependent inhibition of the IFN-α response against subgenomes by lactacystin and epoxomicin. (13A) Cells were incubated with lactacystin and epoxomicin at the indicated concentrations for 7 h and then for an additional 12 h without the drugs. One hour after incubation with the proteasome inhibitors, IFN-α (100 IU/ml) was added for 6 h to a fraction of the cell culture plates (lanes 6 to 10 and 16 to 20). Viral RNA levels were determined by Northern blot analysis. rRNA was used as a control for the amount of RNA present in the samples. (13B and 13C) The amount of viral RNA was measured with a phosphorimager, and values were plotted as percentages of the values obtained with untreated cells.

25

30

35

Figures 14A and 14B are Northern blots and a graph showing that proteasome inhibitors block the IFN- $\alpha$ response against HCV replicons. (14A) GS4.1 cells were left untreated (lanes 1 to 3 and 19 to 21) or treated with 100 IU of IFN- $\alpha$  /ml for 6 h (lanes 4 to 6 and 22 to 24), with 5  $\mu M$  lactacystin (lanes 7 to 9 and 25 to 27) or 1 uM epoxomicin (lanes 13 to 15 and 31 to 33) alone for 7 h, or with 5  $\mu M$  lactacystin (lanes 10 to 12 and 28 to 30) or 1 uM epoxomicin (lanes 16 to 18 and 34 to 36) alone for 1 h and then in the presence of 100 IU of IFN- $\alpha$  /ml 10 for an additional 6 h. Cells were harvested at 12 h (lanes 1 to 18) and 18 h (lanes 19 to 36) after addition of the cytokine. Viral RNA levels were determined by Northern blot analysis. rRNA was used as a control for the amount of RNA present in the samples. (14B) The 15 amount of viral RNA was measured with a phosphorimager and the mean values and standard deviations from three samples were plotted. \*, P < 0.05; \*\*, P < 0.01. PSL, arbitrary units.

20 Figures 15A and 15B show a Northern blot graph demonstrating that proteasome inhibitors prevent establishment of an IFN-α response against HCV replicons. (15A) GS4.1 cells were treated with IFN- $\alpha$  for 10 h, followed by treatment with the indicated proteasome 25 inhibitors for 12 h. Cells were left untreated (lanes 1 to 3 and 4 to 6) or treated with 100 IU of IFN- $\alpha$  /ml for 10 h (lanes 7 to 18), followed by treatment with 10 µM lactacystin (lanes 13 to 15) or 1 µM epoxomicin (lanes 16 to 18) for 12 h. Cells were harvested at 0, 10, and 18 h after the cytokine treatment, as indicated. Viral RNA 30 levels were determined by Northern blot analysis. rRNA was used as a control for the amount of RNA present in the samples. (15B) The amount of viral RNA was measured with a phosphorimager and the mean values and standard 35 deviations from three samples were plotted. PSL,

arbitrary units.

20

25

30

35

### DETAILED DESCRIPTION OF THE INVENTION

The hepatitis C virus (HCV) pandemic affects the health of more than 170 million people and is the major 5 indication for orthotopic liver transplantations (OLT). Although the human liver is the primary site for HCV replication, it is not known whether extrahepatic tissues are also infected by the virus and whether non-primate cells are permissive for RNA replication. 10 because viral replication leads to the accumulation of mutations, it is conceivable that variants can emerge with novel properties such as the potential to replicate in different cell types of various species. Furthermore, accumulation of a large number of quasispecies may also 15 contribute to resistance to IFN- $\alpha$  treatment. Therefore, it is important to determine the properties of HCV variants, and the effect such variation has on the efficacy of IFN- $\alpha$  therapy.

Provided herein is evidence that subgenomic HCV RNAs can replicate in mouse hepatoma and non-hepatic human epithelial cells. Moreover, efficient replication requires adaptation of the virus to cell-type specific environmental conditions. These results show that HCV RNA replication can lead to the accumulation of mutants with altered tissue and host tropism thereby facilitating the development of small animal models for HCV infection.

In accordance with the present invention, there are provided nucleic acids and stably-transfected human non-hepatic, and murine hepatic cell lines that replicate HCV. Also provided are methods of use for such cells for identifying therapeutic anti-viral agents for the treatment of HCV infection. Additionally, the availability of a murine line which replicates HCV enables the production of a greatly needed mouse model of

HCV infection. Furthermore, the invention provides polynucleotides and their corresponding polypeptides which have adaptive mutations which result in expanded tropism of HCV.

The detailed description set forth below describes preferred methods for making and using the nucleic acids and cell lines of the present invention, and for practicing the methods of the invention. Any molecular cloning or recombinant DNA techniques not specifically described are carried out by standard methods, as generally set forth, for example, in Sambrook et al., "DNA Cloning, A Laboratory Manual," Cold Spring Harbor Laboratory, 1989 and Ausubel et al. Current Protocols in Molecular Biology, J. Wiley & Sons, 1995.

15

20

35

10

5

#### I. Definitions

The following definitions are provided to aid in understanding the subject matter regarded as the invention.

As used herein, "hepatitis C virus" or "HCV" shall mean any representative of a diverse group of related viruses classified within the hepacivirus genus of the Flaviviridae family.

"Anti-HCV compounds" may include any inhibitor of

HCV-derived enzymes, such as protease, helicase, and
polymerase inhibitors. Anti-HCV compounds also include
IRES inhibitors, glycosylation inhibitors, and molecules
which block the HCV receptor (thus preventing entry into
cells.) Other anti-HCV compounds include compounds which
enhance the specific or non-specific immune response,
thereby ameliorating HCV infection or symptoms.

"HCV replication levels" may be measured by methods known in the art, including but not limited to detection of replicated HCV replicons, HCV NS protein production, or incorporation of detectably labeled nucleotides into

an HCV nucleic acid.

25

30

35

"Nucleic acid" or a "nucleic acid molecule" as used herein refers to any DNA or RNA molecule, either single or double stranded and, if single stranded, the molecule of its complementary sequence in either linear or circular form. In discussing nucleic acid molecules, a sequence or structure of a particular nucleic acid molecule may be described herein according to the normal convention of providing the sequence in the 5' to 3' direction. With reference to nucleic acids of the 10 invention, the term "isolated nucleic acid" is sometimes This term, when applied to DNA, refers to a DNA molecule that is separated from sequences with which it is immediately contiguous in the naturally occurring genome of the organism or virus in which it originated. 15 When applied to RNA, the term "isolated nucleic acid" refers primarily to an RNA molecule that has been sufficiently separated from other nucleic acids with which it would be associated in its natural state (i.e., in cells or tissues), and explicitly includes viral RNA. 20 An isolated nucleic acid (either DNA or RNA) may further represent a molecule produced directly by biological or synthetic means and separated from other components present during its production.

"RNA subgenome" refers to any molecule which lacks some portion of a genome. For example, an RNA subgenome can be an HCV RNA molecule in which a structural gene has been replaced with a selection agent.

All amino acid residue sequences represented herein conform to the conventional left-to-right amino-terminus to carboxy-terminus orientation.

The term "isolated protein" or "isolated and purified protein" is sometimes used herein. This term refers primarily to a protein produced by expression of an isolated nucleic acid molecule of the invention.

Alternatively, this term may refer to a protein that has been sufficiently separated from other proteins with which it would naturally be associated, so as to exist in "substantially pure" form. "Isolated" is not meant to exclude artificial or synthetic mixtures with other compounds or materials, or the presence of impurities that do not interfere with the fundamental activity, and that may be present, for example, due to incomplete purification, addition of stabilizers, or compounding into, for example, immunogenic preparations or pharmaceutically acceptable preparations.

5

10

"Variants", "mutants" and "derivatives" of particular sequences of nucleic acids refer to nucleic acid sequences that are closely related to a particular sequence but which may possess, either naturally or by 15 design, changes in sequence or structure. By closely related, it is meant that at least about 75%, but often, more than 90%, of the nucleotides of the sequence match over the defined length of the nucleic acid sequence. Changes or differences in nucleotide sequence between 20 closely related nucleic acid sequences may represent nucleotide changes in the sequence that arise during the course of normal replication or duplication in nature of the particular nucleic acid sequence. Other changes may be specifically designed and introduced into the sequence 25 for specific purposes, such as to expand the tropism of viral RNA, or to change an amino acid codon or sequence in a regulatory region of the nucleic acid. specific changes may be made in vitro using a variety of mutagenesis techniques or produced in a host organism 30 placed under particular selection conditions that induce or select for the changes. Such sequence variants generated specifically may be referred to as "mutants" or "derivatives" of the original sequence. The terms "percent similarity", "percent identity" and "percent 35

homology" when referring to a particular sequence are used as set forth in the University of Wisconsin GCG software program.

5

An "adaptive mutation" is a mutation in a nucleic acid sequence which produces a change in viral properties or activity. For example, and adaptive mutation includes, but is not limited to, a mutation which provides enhanced tropism for HCV, or which alters the efficacy of IFN- $\alpha$  treatment.

An HCV peptide, polypeptide, or protein of the 10 invention includes any analogue, fragment, derivative or mutant which is derived from a HCV peptide or polypeptide and which retains at least one property or other characteristic of the HCV polypeptide. Different "variants" of the HCV polypeptide exist in nature. These 15 variants may be alleles characterized by differences in the nucleotide sequences of the gene coding for the protein, or may involve different RNA processing or posttranslational modifications. The skilled person can produce variants having single or multiple amino acid 20 substitutions, deletions, additions or replacements. These variants may include inter alia: (a) variants in which one or more amino acids residues are substituted with conservative or non-conservative amino acids, (b) variants in which one or more amino acids are added to 25 the HCV peptide or polypeptide, (c) variants in which one or more amino acids include a substituent group, and (d) variants in which one or more amino acids are deleted from the HCV peptide or polypeptide. Other HCV peptides or polypeptides of the invention include variants in 30 which amino acid residues from one species are substituted for the corresponding residue in another species, either at the conserved or non-conserved positions. In another embodiment, amino acid residues at non-conserved positions are substituted with conservative 35

5

10

20

25

30

or non-conservative residues. The techniques for obtaining these variants, including genetic (suppressions, deletions, mutations, etc.), chemical, and enzymatic techniques are known to the person having ordinary skill in the art.

To the extent such variations, analogues, fragments, derivatives, mutants, and modifications, including alternative nucleic acid processing forms and alternative post-translational modification forms result in derivatives of the HCV peptide or polypeptide that retain any of the biological properties of the HCV peptide or polypeptide, they are included within the scope of this invention.

The term "functional" as used herein implies that

15 the nucleic or amino acid sequence is functional for the
recited assay or purpose.

The phrase "consisting essentially of" when referring to a particular nucleotide or amino acid means a sequence having the properties of a given sequence. For example, when used in reference to an amino acid sequence, the phrase includes the sequence per se and molecular modifications that would not affect the fundamental and novel characteristics of the sequence.

A "replicon" is any genetic element, for example, a plasmid, cosmid, bacmid, phage or virus, that is capable of replication largely under its own control. A replicon may be either RNA or DNA and may be single or double stranded.

A "vector" is a replicon, such as a plasmid, cosmid, bacmid, phage or virus, to which another genetic sequence or element (either DNA or RNA) may be attached so as to bring about the replication of the attached sequence or element.

The phrase "operably linked" when referring to nucleic acid constructs is used herein to indicate that

5

10

the respective promoter, operator and coding sequences, as well as any other 5' and 3' regulatory sequences, are arranged in the appropriate location, order and reading frame such that the desired control (e.g., expression) is effected under appropriate conditions.

The term "oligonucleotide," as used herein refers to primers and probes of the present invention, and is defined as a nucleic acid molecule comprised of two or more ribo- or deoxyribonucleotides, preferably more than three. The exact size of the oligonucleotide will depend on various factors and on the particular application and use of the oligonucleotide.

The term "probe" as used herein refers to an oligonucleotide, polynucleotide or nucleic acid, either RNA or DNA, whether occurring naturally as in a purified 15 restriction enzyme digest or produced synthetically, which is capable of annealing with or specifically hybridizing to a nucleic acid with sequences complementary to the probe. A probe may be either single-stranded or double-stranded. The exact length of 20 the probe will depend upon many factors, including temperature, source of probe and use of the method. For example, for diagnostic applications, depending on the complexity of the target sequence, the oligonucleotide probe typically contains 15-25 or more nucleotides, 25 although it may contain fewer nucleotides. The probes herein are selected to be "substantially" complementary to different strands of a particular target nucleic acid sequence. This means that the probes must be sufficiently complementary so as to be able to 30 "specifically hybridize" or anneal with their respective target strands under a set of pre-determined conditions. Therefore, the probe sequence need not reflect the exact complementary sequence of the target. For example, a non-complementary nucleotide fragment may be attached to 35

the 5' or 3' end of the probe, with the remainder of the probe sequence being complementary to the target strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the sequence of the target nucleic acid to anneal therewith specifically.

5

35

The term "specifically hybridize" refers to the association between two single-stranded nucleic acid molecules of sufficiently complementary sequence to 10 permit such hybridization under pre-determined conditions generally used in the art (sometimes termed "substantially complementary"). In particular, the term refers to hybridization of an oligonucleotide with a substantially complementary sequence contained within a 15 single-stranded DNA or RNA molecule of the invention, to the substantial exclusion of hybridization of the oligonucleotide with single-stranded nucleic acids of non-complementary sequence. For example, hybridizations may be performed, according to the method of Sambrook et 20 al., Molecular Cloning, Cold Spring Harbor Laboratory (1989), using a hybridization solution comprising: 5X SSC, 5X Denhardt's reagent, 1.0% SDS, 100  $\mu$ g/ml denatured, fragmented salmon sperm DNA, 0.05% sodium pyrophosphate and up to 50% formamide. Hybridization is 25 carried out at 37-42°C for at least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1% SDS; (3) 30 minutes-1 hour at 37°C in 1X SSC and 1% SDS; (4) 2 30 hours at 42-65°C in 1X SSC and 1% SDS, changing the solution every 30 minutes.

One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence homology

(Sambrook et al., 1989) is as follows:

 $T_m = 81.5$ °C + 16.6Log [Na+] + 0.41(% G+C) - 0.63 (% formamide) - 600/#bp in duplex

5

10

As an illustration of the above formula, using [Na<sup>+</sup>] = [0.368] and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the  $T_m$  is 57°C. The  $T_m$  of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C.

The stringency of the hybridization and wash depend primarily on the salt concentration and temperature of the solutions. In general, to maximize the rate of 15 annealing of the probe with its target, the hybridization is usually carried out at salt and temperature conditions that are 20-25°C below the calculated  $T_{m}$  of the hybrid. Wash conditions should be as stringent as possible for the degree of identity of the probe for the target. In 20 general, wash conditions are selected to be approximately 12-20°C below the  $T_{m}\ \text{of the hybrid.}$  In regards to the nucleic acids of the current invention, a moderate stringency hybridization is defined as hybridization in 6 X SSC, 5 X Denhardt's solution, 0.5% SDS and 100  $\mu \text{g/ml}$ 25 denatured salmon sperm DNA at 42°C, and washed in 2X SSC and 0.5% SDS at 55°C for 15 minutes. A high stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100  $\mu$ g/ml denatured salmon sperm DNA at 42°C, and washed in 1% SSC and 0.5% 30 SDS at 65°C for 15 minutes. A very high stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100  $\mu$ g/ml denatured salmon sperm DNA at 42°C, and washed in 0.1% SSC and 0.5% SDS at 65°C for 15 minutes. 35

The term "primer" as used herein refers to an oligonucleotide, either RNA or DNA, either singlestranded or double-stranded, either derived from a biological system, generated by restriction enzyme digestion, or produced synthetically which, when placed in the proper environment, is able to functionally act as an initiator of template-dependent nucleic acid synthesis. When presented with an appropriate nucleic acid template, suitable nucleoside triphosphate precursors of nucleic acids, a polymerase enzyme, suitable cofactors and conditions such as a suitable temperature and pH, the primer may be extended at its 3' terminus by the addition of nucleotides by the action of a polymerase or similar activity to yield a primer 15 extension product. The primer may vary in length depending on the particular conditions and requirement of the application. For example, in diagnostic applications, the oligonucleotide primer is typically 15-25 or more nucleotides in length. The primer must be of sufficient complementarity to the desired template to 20 prime the synthesis of the desired extension product, that is, to be able anneal with the desired template strand in a manner sufficient to provide the 3' hydroxyl moiety of the primer in appropriate juxtaposition for use in the initiation of synthesis by a polymerase or similar 25 enzyme. It is not required that the primer sequence represent an exact complement of the desired template. For example, a non-complementary nucleotide sequence may be attached to the 5' end of an otherwise complementary primer. Alternatively, non-complementary bases may be 30 interspersed within the oligonucleotide primer sequence, provided that the primer sequence has sufficient complementarity with the sequence of the desired template strand to functionally provide a template-primer complex for the synthesis of the extension product. 35

As used herein, the terms "reporter," "reporter system", "reporter gene," or "reporter gene product" shall mean an operative genetic system in which a nucleic acid comprises a gene that encodes a product that when expressed produces a reporter signal that is a readily measurable, e.g., by biological assay, immunoassay, radioimmunoassay, or by colorimetric, fluorogenic, chemiluminescent or other methods. The nucleic acid may be either RNA or DNA, linear or circular, single or double stranded, antisense or sense polarity, and is operatively linked to the necessary control elements for the expression of the reporter gene product. required control elements will vary according to the nature of the reporter system and whether the reporter gene is in the form of DNA or RNA, but may include, but not be limited to, such elements as promoters, enhancers, translational control sequences, poly A addition signals, transcriptional termination signals and the like.

10

15

20

25

30

35

The terms "transform", "transfect", "transduce", shall refer to any method or means by which a nucleic acid is introduced into a cell or host organism and may be used interchangeably to convey the same meaning. Such methods include, but are not limited to, electroporation, microinjection, PEG-fusion and the like.

The introduced nucleic acid may or may not be integrated (covalently linked) into nucleic acid of the recipient cell or organism. In bacterial, yeast, plant and mammalian cells, for example, the introduced nucleic acid may be maintained as an episomal element or independent replicon such as a plasmid. Alternatively, the introduced nucleic acid may become integrated into the nucleic acid of the recipient cell or organism and be stably maintained in that cell or organism and further passed on or inherited to progeny cells or organisms of the recipient cell or organism. In other applications,

the introduced nucleic acid may exist in the recipient cell or host organism only transiently.

A "clone" or "clonal cell population" is a population of cells derived from a single cell or common ancestor by mitosis.

5

15

20

25

30

35

A "cell line" is a clone of a primary cell or cell population that is capable of stable growth in vitro for many generations.

A "selectable marker" or a "selection agent" refers
to a nucleic acid sequence that when expressed confers a
selectable phenotype, such as antibiotic resistance, to a
transformed cell.

A "viral antigen" shall be any peptide, polypeptide or protein sequence, segment or epitope that is derived from a virus that has the potential to cause a functioning immune system of a host to react to said viral antigen.

An "antibody" or "antibody molecule" is any immunoglobulin, including antibodies and fragments thereof, that binds to a specific antigen. The term includes polyclonal, monoclonal, chimeric, and bispecific antibodies. As used herein, antibody or antibody molecule contemplates both an intact immunoglobulin molecule and an immunologically active portion of an immunologically molecule such as those portions known in the art as Fab, Fab', F(ab')2 and F(v).

The term "detectably label" is used herein to refer to any substance whose detection or measurement, either directly or indirectly, by physical or chemical means, is indicative of the presence of the target bioentity in the test sample. Representative examples of useful detectable labels, include, but are not limited to the following: molecules or ions directly or indirectly detectable based on light absorbance, fluorescence, reflectance, light scatter, phosphorescence, or

luminescence properties; molecules or ions detectable by their radioactive properties; molecules or ions detectable by their nuclear magnetic resonance or paramagnetic properties. Included among the group of molecules indirectly detectable based on light absorbance or fluorescence, for example, are various enzymes which cause appropriate substrates to convert, e.g., from non-light absorbing to light absorbing molecules, or from non-fluorescent to fluorescent molecules.

As used herein, the term "living host" shall mean any non-human autonomous being.

II. Methods for Obtaining HCV RNA and Producing Non-Hepatic Human Cell Lines and Non-Human Hepatic Cell Lines that Replicate HCV.

The HCV replicating non-hepatic human cell-based and non-human hepatic cell-based systems are prepared according to the general methods set forth below for isolation of nucleic acids, transformation of cultured cells, and maintenance of cell lines.

### A. Nucleic acids

5

10

15

20

25

30

35

The HCV replicons of the present invention comprise adaptive mutations which alter the ability of HCV to replicate in different cell types. Surprisingly, the present inventors have identified mutations which are associated with expanded viral tropism.

The HCV nucleic acid molecules of the invention may be prepared by two general methods: (1) They may be synthesized from appropriate chemical starting materials, or (2) they may be isolated from biological sources. Both methods utilize protocols well known in the art.

The availability of nucleotide sequence information enables preparation of an isolated nucleic acid molecule of the invention by oligonucleotide synthesis. Synthetic

oligonucleotides may be prepared by the phosphoramadite method employed in the Applied Biosystems 38A DNA Synthesizer or similar devices. The resultant construct may be purified according to methods known in the art, such as high performance liquid chromatography (HPLC). Long, double-stranded polynucleotides, such as a DNA molecule of the present invention, must be synthesized in stages due to the size limitations inherent in current oligonucleotide synthetic methods. Thus, for example, a 3 kilobase double-stranded molecule may be synthesized as 10 several smaller segments of appropriate complementarity. Complementary segments thus produced may be ligated such that each segment possesses appropriate cohesive termini for attachment of an adjacent segment. Adjacent segments may be ligated by annealing cohesive termini in the 15 presence of DNA ligase to construct an entire 3 kilobase double-stranded molecule. A synthetic DNA molecule so constructed may then be cloned and amplified in an appropriate vector.

20 HCV nucleic acid sequences may be isolated from appropriate biological sources using methods known in the art. For example, total RNA can be extracted with TRIzol reagent from Gibco BRL, although other reagents are also available for this purpose.

In some cases, it may be desirable to synthesize HCV subgenomic RNA wherein a selectable marker gene is substituted for a HCV structural gene.

25

30

35

The availability of HCV replicon encoding nucleic acids enables the production of strains of laboratory mice carrying part or all of the HCV sequence or mutated sequences thereof. Such mice provide an *in vivo* model for studying HCV infection, and analyzing possible treatment modalities for the same.

Methods of introducing transgenes in laboratory mice are known to those of skill in the art. Three common

methods include: 1. integration of retroviral vectors encoding the foreign gene of interest into an early embryo; 2. injection of DNA into the pronucleus of a newly fertilized egg; and 3. the incorporation of genetically manipulated embryonic stem cells into an early embryo.

A transgenic mouse carrying an HCV replicon comprising the adaptive mutations is generated by genomic integration of exogenous genomic sequence encoding HCV. These transgenic animals are useful for drug screening studies as animal models for human diseases.

10

35

The term "animal" is used herein to include all vertebrate animals, except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages. A "transgenic animal" is any 15 animal containing one or more cells bearing genetic information altered or received, directly or indirectly, by deliberate genetic manipulation at the subcellular level, such as by targeted recombination or microinjection or infection with recombinant virus. The 20 term "transgenic animal" is not meant to encompass classical cross-breeding or in vitro fertilization, but rather is meant to encompass animals in which one or more cells are altered by or receive a recombinant DNA molecule. This molecule may be specifically targeted to 25 a defined genetic locus, be randomly integrated within a chromosome, or it may be extrachromosomally replicating The term "germ cell line transgenic animal" refers to a transgenic animal in which the genetic alteration or genetic information was introduced into a germ line cell, 30 thereby conferring the ability to transfer the genetic information to offspring. If such offspring, in fact, possess some or all of that alteration or genetic information, then they, too, are transgenic animals.

A type of target cell for transgene introduction is the embryonal stem cell (ES). ES cells may be obtained from pre-implantation embryos cultured in vitro (Evans et

al., (1981) Nature 292:154-156; Bradley et al., (1984) Nature 309:255-258; Gossler et al., (1986) Proc. Natl. Acad. Sci. 83:9065-9069). Transgenes can be efficiently introduced into the ES cells by standard techniques such as DNA transfection or by retrovirus-mediated transduction. The resultant transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The introduced ES cells thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal.

#### B. Cell lines

5

10

15

20

The cell lines of the invention include any cell which supports production of HCV components. These cells include human, non hepatic cells and/or non-human hepatic cells, such as murine hepatic cells. Cell lines useful for practice of the invention include, but are not limited to HELA, a non-hepatic epithelial cell line (ATCC CRL number CCL-2.2), Hepa1-6, a murine hepatoma cell line (ATCC CRL number-1830), and AML-12, a murine hepatocyte cell line (ATCC CRL number-2254).

To achieve stable gene transfer, HCV subgenomic RNA

is introduced into host cells. This may be accomplished according to numerous methods known in the art, including, but not limited to: (1) calcium phosphate 25 transfection; (2) transfection with DEAE-dextran; (3) electroporation; and (4) liposome-mediated transfection. For general protocols, see, e.g., chapter 9 in Current Protocols in Molecular Biology, Ausubel et al. (editors), John Wiley & Sons, Inc. 1987-1995. For stable transfer of 30 nucleic acids into mammalian cells, the liposome-mediated transfection method may be used in the present invention because of the large amount of nucleic acid that can be introduced into the cells, thereby increasing the possibility of integration of the nucleic acid into the 35 host genome.

5

35

Cells are grown according to standard methods known in the art, such as those set forth in <u>Culture of Animal Cells: A Manual of Basic Technique</u> by R. Ian Freshney, 4th ed. Edition, available from the ATCC.

Stable transfectants are selected by the ability of an individual cell colony to grow in the presence of a selection agent, e.g., an antibiotic, by virtue of a resistance-encoding gene on HCV RNA or by isolating cells using FACS and antibodies directed any HCV protein.

Detection and quantitation of expression of HCV gene 10 products in stably-transfected cell lines of the invention can be accomplished using a variety of known assays. For instance, cells transformed with the RNA subgenomes of HCV can be selected with an antibiotic such as G418 (neomycin). Alternatively, cells may be selected 15 based on the presence and accumulation of HCV RNA or HCV gene products. As another example, the starting HCV encoding nucleic acids may be modified to also comprise a hygromycin or puromycin or any other resistance or reporter gene, such that cells transfected with the 20 nucleic acids can be selected by their ability to grow on hygromycin- or puromycin-containing medium. Alternatively, selectable markers including luciferase, beta lactamase etc., may be utilized which allow for the selection of cells by FACS and related procedures. In an 25 alternative embodiment, a separate plasmid may be constructed that comprises an antibiotic resistance gene, and can be used to co-transfect cells along with the subgenomic RNA molecules. Further, as described in detail in the following Examples, cells stably 30 transfected with the subgenomic HCV RNA are grown in the appropriate medium for a selected period of time, the medium is then collected and analyzed for the presence of HCV RNA by dot blot hybridization or by conventional

Northern hybridization, using a radioactively labeled

5

35

probe having HCV DNA or RNA complementary sequences. Alternatively, viral gene products may be detected in the cells of the invention using conventional methods, including, without limitation, immunoassay and Western blotting.

Using the assays described above, stably-transfected cell lines can be selected which possess optimum characteristics for use in cell-based assays for screening potential anti-viral compounds.

Another aspect of the invention includes a non-human host animal which comprises the HCV expressing cells of the invention. These animals may be produced by administration of a HCV replicating cell, an HCV encoding nucleic acid having one or more adaptive mutations which permit replication in mice. The cells or viral nucleic acid could be directly injected intravenously (e.g. via tail vein injection), intramuscularly, subcutaneously, or via-intrahepatic injection. Alternatively, transgenic mice could be produced using the HCV replicons of the inventions, as described above.

# III. Uses of Cell Lines for Cell-Based Assays of Potential Anti-HCV Agents

The human non-hepatic and murine hepatic cell lines

of the invention which replicate HCV may be used in
research, diagnostic, and therapeutic applications,
including cell-based assays to evaluate the effectiveness
of potential anti-HCV compounds, utilizing methodologies
known in the art. Typical assays are summarized herein

below. These cell-based assays may be performed in
standard cell culture media utilizing commonly-available
equipment, reagents and culture containers.

Persons skilled in the art will appreciate that these assays represent exemplary embodiments, and may be varied to provide similar/equivalent equipment or

reaction conditions. For example, a variety of genes encoding antibiotic resistance are available, and can be utilized in accordance with the present invention in the generation of the cell lines of the invention. preferred embodiment, RNA isolated from parental human ٠ 5 hepatic or untransformed cells is also utilized as a control in the assays described herein below to determine the effects of potential anti-viral compounds on HCV expressed in the cells. The control RNA is obtained in a manner similar to the HCV RNA. This cell line is treated 10 in the assays described herein below as a negative control, to assure that any effects observed are due to the action of the compound being tested on HCV, and not non-specific effects due to the introduction of RNA into 15 the cells.

# A. General Cell-Based Assay for Inhibitors of HCV replication

96-well microtiter plates are seeded with an appropriate amount of cells which replicate HCV in a 20 standard cell culture medium containing G418 (e.g., 400 µg/ml), as well as standard concentrations of penicillin, streptomycin and kanamycin or gentamicin to prevent bacterial and mycoplasma contamination. The cells are 25 incubated at 37°C in a humidified 5% CO2 incubator. On day O wells are washed three times with warm phosphatebuffered saline (PBS). The culture medium is then replaced with fresh medium containing 0.3% dimethylsulfoxide (DMSO), 10% fetal calf serum (FCS), penicillin, streptomycin, kanamycin/gentamicin, 30 containing one of the following ingredients: (1) various concentrations of a known HCV inhibitor, such as interferon alpha, as a positive control; and (2) various concentrations of one or more of the compounds to be The plates are incubated at 37°C in humidified, 35 tested.

5% CO2 incubator for 24, 48, and 72 hours. The plates are washed twice with PBS and then with a solution of methanol and acetone (1:1) to fix the cells. are then incubated with an antibody specific for a viral protein (i.e. NS5A) according to the standard methods, such as enzyme linked immunosorbent assay (ELISA). Briefly, following incubation with the primary antibody, the plates are washed to remove unbound antibody and then incubated with a second, enzyme-conjugated antibody that can bind to the primary antibody. The plates are washed 10 again, followed by an incubation with a colorless substrate that upon hydrolysis (cleavage) by the enzyme yields a colored product, the concentration of which can be determined with a spectrophotometer (microtiter plate reader). The concentration of the product corresponds to 15 the levels of viral replication in cells and can be used to determine the activity of a given drug to inhibit HCV replication.

### 20 B. Cytotoxicity Assays

25

30

A cytotoxicity assay may be conducted to evaluate potential anti-HCV agents, utilizing a protocol similar to that described above. Instead of measuring HCV replication levels, however, cytoxicity of the various test agents is assessed by standard procedures to determine cell viability, proliferation and levels of cellular metabolism including but not restricted to cell membrane permeability, lysosomal mass-pH, cell density or mitochondrial activity. For example, the CytoTox-ONE<sup>TM</sup> Assay from Promega is a rapid, fluorescent measure of the release of lactate dehydrogenase (LDH) from cells with a damaged membrane. LDH released into the culture medium is measured with a 10-minute coupled enzymatic assay that results in the conversion of resazurin into resorufin.

35 Since the CytoTox-ONE™ Reagent mix does not damage

healthy cells, released LDH can be measured directly in assay wells containing a mixed population of viable and damaged cells.

# 5 IV. Identification of Cell Lines Permissive for HCV Infection

As shown herein, it is possible to produce HCV carrying adaptive mutations that confer broad tissue and species tropism. Using such virus stocks it will be possible to screen cell lines of human and non-human origin for virus infection. Briefly, probes which correspond to unique portions of the sequence, may be used in detection methods. This method will lead to the identification of novel cell lines that are permissive for a complete cycle of HCV replication.

### V. Screening for the HCV Receptor(s)

10

15

The ability to replicate HCV in different cell lines facilitates the isolation of the HCV receptor. Virus stocks similar to the ones described in section IV can be 20 used to isolate the HCV receptor(s). For this purpose virus stocks carrying replicons with a selectable marker, such as neomycin or hygromycin will be used. Cells that are non-permissive for infection, will be transfected with DNA isolated from cells that are known to express 25 the receptor (i.e. human hepatocytes, cells identified with the procedure described in section III) and subsequently infected with recombinant HCV carrying the selectable marker. Cells that express the receptor can then be selected through the addition of an antibiotic 30 (i.e. G418 or hygromycin) to the culture medium. cells are identified, the transfected DNA can be isolated, cloned, and sequenced. The sequence information can then be used to identify the gene(s) encoded by transfected DNA. 35

The following examples are provided to describe the invention in further detail. These examples are intended to illustrate and not to limit the invention.

5

35

# EXAMPLE I Human Non-Hepatic and Mouse Hepatic Cell Lines that replicate HCV

#### MATERIALS AND METHODS

- Cell culture. Cells were purchased from the American Type Culture Collection (BHK Kidney Mesocricetus auratus (Syrian golden hamster) ATCC CRL-1632; Vero Kidney epithelial Cercopithecus aethiops (African green monkey) ATCC CCL-81; CV-1 Kidney fibroblast Cercopithecus

  aethiops (African green monkey) ATCC CCL-70; HT1080
  Fibrosarcoma Homo sapiens (human) ATCC CRL12012; HeLa Cervix carcinoma Homo sapiens (human) ATCC CRL12012; McA-RH7777 Hepatoma Rattus norvegicus (rat) ATCC CRL-1601;
  FT02B Hepatoma Rattus norvegicus (rat); Hepa1-6 Hepatoma

  Mus musculus (mouse) ATCC CRL-2254; FL83B Hepatocyte Mus
- Mus musculus (mouse) ATCC CRL-1830; AML12 Hepatocyte Mus musculus (mouse) ATCC CRL-2254; FL83B Hepatocyte Mus musculus (mouse) ATCC CRL-2390). The Huh7-derived cell lines GS4.1 and GS4.5 are subclones derived from cell lines FCA1 and FCA4, respectively (Guo, J. T., et al.,
- 25 2001., J. Virol. 75:8516-8523). Cell line Bsp8 is a Huh7-derived cell line expressing HCV-N subgenomic replicon 1bneoΔS (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). All cultures were grown in Dulbecco's modified Eagle's medium (Gibco-Invitrogen) supplemented with 10%
- 30 fetal bovine serum, L-glutamine, nonessential amino acids, penicillin, and streptomycin.
  - RNA transfection. All the plasmids were linearized with ScaI, and RNA was synthesized with the MEGAscript kit (Ambion). In vitro-transcribed RNA was purified as

previously described (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). Total cellular RNA was extracted with Trizol reagent (Invitrogen). The conditions used for the transfection of cells with total RNA were identical to those used for the transfection with in vitrotranscribed RNA (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). Colonies were selected with G418 at a concentration of 1 mg/ml.

10 RNA analysis. Total cellular RNA was extracted from transfected cell lines with Trizol reagent. Five micrograms of total RNA was fractionated on 1% agarose gels containing 2.2 M formaldehyde and transferred onto a nylon membrane. Membranes were hybridized with riboprobes specific for plus-stranded HCV replicon RNA, human papillomavirus (HPV) E6, and mouse albumin mRNA as described previously (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). The HPV and mouse albumin probes spanned nucleotides 811 to 1491 (GenBank accession number M20325) and nucleotides 1501 to 1988 (GenBank accession number XM\_132149), respectively.

Reverse transcription-PCR and DNA sequencing. Nucleotide and amino acid numbers correspond to the HCV type 1b genome Con-1 (AJ238799). HCV replicons were isolated and 25 cloned from established cell lines by PCR amplification of three fragments spanning the entire NS region from position 3420 to 9410. The untranslated regions at the 5' and 3' ends of HCV RNA were cloned separately for nucleotide sequence analysis. DNA synthesis was carried 30 out with Superscript II reverse transcriptase provided in a cDNA synthesis kit (Gibco- Invitrogen). The DNA oligomers used as primers for the reverse transcription reaction mapped to positions 485 to 465, 5492 to 5473, 7256 to 7234, 9410 to 9388, and 9616 to 9597. The 35

reaction mixtures were incubated for 1 h at 45°C. PCR was performed with an Advantage PCR kit (Clontech). One microliter of the cDNA reaction mixture was used for PCRs with 19- to 23-nucleotide-long primers that yielded

5 fragments spanning positions 1 to 464, 1387E to 5082, 5016 to 7226, 7154 to 9387, and 9239 to 9616. Position 1387E refers to an oligomer specific for the encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) element located upstream of NS3. The PCR products were cloned into plasmid pGEM-T Easy (Promega). Four clones of each fragment were sequenced with an ABI automatic DNA sequencer, and a consensus sequence was established with the help of a sequence assembly program (Genetics Computer Group).

Long reverse transcription-PCR was performed with an Advantage-GC kit (Clontech) with a pair of primers beginning at positions 1415E, upstream of NS3, and 7989 within NS5B. The PCR conditions were modified as follows: step 1, 95°C for 3 min; step 2, 5 cycles, 30 s at 95°C and 6 min at 72°C; step 3, 27 cycles, 30 s at 95°C and 6 min at 68°C; step 4, 68°C for 6 min. PCR products were gel purified and digested with HindIII and MfeI and replaced with the corresponding fragment in plasmid 1377/NS3-3'.

25

30

Plasmid construction. All plasmids (Table 3) were derived from the parental HCV Con-1 replicon I377/NS3-3' (AJ242652). Subgenomes containing consensus mutations were constructed by replacing DNA restriction fragments with the corresponding fragments from the pGEM-T Easy cDNA libraries (see above). The resulting plasmids with the amino acid changes in the NS region are listed in Table 3.

35 Immunofluorescence. Cells were plated on coverslips in

PCT/US2003/039722 WO 2004/055216

six-well plates at least 16 h before treatment, washed with phosphate-buffered saline, and fixed with cold methanol-acetone (1:1) for 15 to 20 min. Next, the cells were blocked in phosphate-buffered saline containing 10% fetal bovine serum for 30 min at room temperature and then incubated with anti-NS5A antibodies (a gift from Chen Liu) and fluorescein isothiocyanate-conjugated goat anti-mouse immunoglobulin antibodies (Jackson Laboratories). In addition, cells were stained with the DNA binding fluorochrome DAPI (4',6'-diamidino-2-10 phenylindole). Coverslips were mounted with antifade agent (Molecular Probes), examined with a Nikon immunofluorescence microscope, and photographed with a charge-coupled device camera.

15

35

5

#### RESULTS

### HCV replication in cells of nonhepatic origin.

HCV exhibits a very narrow host range and infects only humans and chimpanzees. We question whether this limitation was due to determinants of RNA replication. 20 Because efficient replication of subgenomes depends on genetic adaptations of the replicon (Blight, K. J., et al. 2000. Science 290:1972-1975; Guo, J. T., et al., 2001., J. Virol. 75:8516-8523; Lohmann, V., et al., 2001. J. Virol. 75:1437-1449), presumably to compensate for 25 subtle variations in the cellular environments among cells from different tissues, it was hypothesized that replication in cells of nonhepatic origin would require additional, cell-type-specific adaptive mutations. Transfection of several primate- and rodent-derived cell 30 lines with subgenomic RNA transcribed from plasmid DNA carrying previously identified adaptive mutations in Huh7 cells did not yield cell lines expressing replicons.

increase the chance for the selection of RNA subgenomes

capable of replicating in cells of nonhepatic origin,

subgenomic RNA isolated from Huh7 cell lines that replicate HCV RNA was used. Because of the high rate of nucleotide incorporation errors that occur during RNA-directed RNA synthesis, this population of viral subgenomes exhibited much greater genetic heterogeneity than did RNA transcribed from a DNA template in vitro.

5

Upon transfection of HeLa cells with total RNA obtained from Huh7 cell lines GS4.1, GS4.5, and Bsp8, G418-resistant cell clones were obtained. The number of clones ranged from approximately 2 (Bsp8) to 50 (GS4.1) 10 per 10  $\mu$ g of total RNA depending on the origin of the RNA used for the transfections. Replicons in these three Huh7-derived cell lines contained different adaptive mutations and replicated two different HCV 1b genomes (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). 15 Several HeLa-derived colonies obtained with total RNA from GS4.1 cells were subsequently expanded into seven stable cell lines (SL1 to SL7; Fig. 1A, lanes 4 and 8 to 12). The amounts of viral RNA present in early passages of these cell lines examined ranged from 0.05 to 7.5 20 ng/10 µg of total RNA, which corresponded to 20 to 3,000 copies of RNA per cell. In general, the amounts of RNA increased upon passage of cells and reached levels that were comparable to those obtained with the most productive Huh7- derived cell lines such as GS4.1 (lanes 25 2 and 4 to 6). As expected, expression of viral gene products could be confirmed by immunofluorescence with antibodies directed against NS5A (Fig. 1B). As with GS4.1 cells, more than 90% of SL1 cells expressed viral proteins. However, in contrast to Huh7 cell lines where 30 the accumulation of HCV RNA declines approximately 100fold when cells become confluent, viral replication in HeLa cells was not affected by the growth conditions of the cells, i.e., SL1 cells continued to produce high amounts of viral RNA even when they became confluent 35

(results not shown) (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523; Pietschmann, T., et al. 2001. J. Virol. 75:1252-1264).

### 5 Adaptation of HCV replicons.

10

15

20

25

30

35

To determine whether HCV replication in HeLa cells led to the selection of subgenomes with cell-typespecific adaptive mutations, the efficiency by which G418-resistant colonies formed in Huh7 and HeLa cells transfected with total RNA isolated from GS4.1 and SL1 cells was compared. Total RNA from GS4.1 cells led to the selection of approximately 166 G418-resistant colonies per ng of viral RNA in Huh7 cells compared with only 4 colonies in HeLa cells (Table 1). In contrast, total RNA from SL1 cells yielded 160 colonies in HeLa cells compared with about 20 in Huh7 cells. These results indicated that replication in HeLa cells led to the selection of variants with cell-type-specific adaptive mutations that were responsible for the 40-fold increase in colony formation efficiency between amplified RNA in GS4.1 and SL1 cells. Nucleotide sequence analysis of HCV cDNA clones obtained from the SL1 and SL2 cell lines confirmed this view. These data showed that replicons in the two HeLa cell lines maintained the previously identified adaptive mutations in GS4.1 cells and acquired several additional mutations that resulted in amino acid changes in the NS region (Figure 2 and Table 2). Notably, some of the new mutations formed clusters in the NS4B and NS5A regions. In the case of SL1 cells; a deletion of 43 amino acids near the C terminus of NS5A was observed. Of particular interest were mutations in the amino-terminal region of NS4B, because they have so far not been found in cDNAs from replicons in Huh7 cells and hence could have been responsible for the observed adaptation of replicating RNA (Blight, K. J., et al. 2000. Science

290:1972-1975; Guo, J. T., et al., 2001., J. Virol. 75:8516-8523; Krieger, N., et al., 2001. J. Virol. 75:4614-4624; Lohmann, V., et al., 2001. J. Virol. 75:1437-1449). Moreover, one mutation at position 1749 was present in both SL1 and SL2 cells. In contrast to the results obtained with the NS regions, no mutations were detected in the 5' and 3' untranslated regions of replicons expressed in SL1 and SL2 cells.

Table 1. Colony formation efficiency of total cellular  $RNA^a$ .

15		N	o. of	colo	nies in tra	nsfect	ed c	ells			
		•	Huh7		HeLa			Hepa1-6			
20	Cell Vi:	ral RNA g/10 μg)	Mean	(SD)	Colonies/ ng of viral RNA	Mean	(SD)	Colonies/ ng of viral RNA	Mean	(SD)	Colonies/ ng of viral RNA
	GS4.1 5	5	834	(64)	166	22	(4)	4	0		<1
	SL1 5	5	100	(53)	20	80	3 (81	.) 160	1.3	(1.	5) <1
25	MH1 0.	5	20	(2)	40	66	(9)	132	1.7	(0.	6) 3
	a Results	s from t	hree i	ndepe	endent tran	sfecti	on ex	cperiments.	Total	RNA	was
	extracte	d from G	S4.1,	SL1,	and MH1 ce	lls at	: pas	sages 21, 2	26, and	14,	
	respecti	vely.									

TABLE 2. Consensus mutations in replicons isolated from HeLa and mouse hepatoma cell clones

Cell clone	Conserved mutation(s)	NS protei
GS4.1	E1202G	NS3
	S2204I, D2254E, I2324V	NS5A
SL1	Q1067R, S1128A, E1202G, S1323P, S1560Ga	NS3
	L1701F	NS4A
	Q1720R, Q1727H, V1749A, V1893L	NS4B
	T2035A, S2204I, I2252V, D2254E, I2274V, R22	90L,
	I2324V, del.2371-2413 <sup>b</sup>	NS5A
•	w2990R	NS5B
SL2	I1097V, Q1112R, P1115L, V1593M, M1647I	NS3
	L1715P, Q1737R, V1749A, I1797V, N1965Y	NS4B
	Q2012L, S2204I, E2247G, D2254E, K2302R, I23	24V,
	S2336P, L2400S, E2411G, A2412V	NS5A
MH1	Q1067R, S1128A, E1202G, S1323P, S1560G	NS3
	Q1720R, Q1727H, V1749A, V1893L	NS4B
	T2035A, S2204I, I2252V, D2254E, I2274V, R22	90L,
	I2324V, M2388T, T2496A	NS5A
	w2990R	NS5B
MH2	Q1112R, E1202G, a S1323P, S1560G	NS3
	L1701F	NS4A
	Q1720R, Q1727H, V1749A, V1893L	NS4B
	T2035A, T2185A, S2204I, I2252V, D2254E, I22	74V,
	R2290 <sup>a</sup> , I2324V <sup>a</sup>	NS5A
	W2990R <sup>a</sup>	NS5B
мн4	Q1067R, S1128A, E1202G, S1323P, S1560G	NS3
	L1701F	NS4A
	Q1720R, Q1727H, V1749A, V1893L, A1841T	NS4B
	T2035A, S2204I, I2252V, D2254E, I2274V, R22	90L,
	I2324V, T2364M, L2391R	NS5A
	I2843V, W2990R	NS5B

a Mutation that occurred in 50% of the clones analyzed.

### Mouse hepatoma cells can support HCV RNA replication.

40 The discovery of several additional mutations in cDNA clones obtained from SL1 and SL2 cells suggested total RNA from these cell lines might yield colonies in cells that did not appear to be permissive for HCV replication after transfection with subgenomic RNA or total RNA from Huh7-derived cell lines. Hepatoma and hepatocyte-derived cell lines were examined. G418-

b del., deletion.

resistant colonies were obtained with the mouse hepatoma cell line Hepa1-6 after transfection with total RNA from SL1 cells (Figure 3A, lanes 4 to 6, 9, and 12). As with HeLa cells, the amounts of RNA ranged from 300 to 1,000 copies of RNA per cell and a large fraction of the cells expressed viral proteins (Figure 3B). In contrast to Huh7 and HeLa cells, the amount of HCV RNA in the mouse cell lines appeared to vary between cell passages (Figure 3A, lanes 6 to 14). Interestingly, total RNA isolated from one of the mouse cell lines, MH1, did not produce 10 significantly more colonies in Hepa1-6 cells than did total RNA from SL1 cells, suggesting that the subgenomes present in SL1 cells were already adapted for replication in the mouse cells (Table 1). In support of this interpretation, nucleotide sequence analysis of viral 15 cDNAs cloned from three mouse cell lines showed that the majority of the mutations identified in SL1 cells were maintained (Figure 2). Surprisingly, the deletion in NS5A identified in four of four clones sequenced from SL1 cells was not present in replicons isolated from mouse 20 cells, indicating that a subpopulation of replicons without the deletion was still present in these (SL1) cells.

In further experiments, mouse hepatocyte cells AML12

(ATCC CRL-2254) were transfected with total RNA isolated from the cell line GS4.1, expressing subgenomic replicons and from cell line HFL expressing full-length HCV genomes, respectively. G418 resistant colonies were isolated to establish stable cell lines expressing HCV suggenomic and full-length replicons. 5 micrograms of total RNA was isolated from AML12 cell lines (MA6-5 to MA6-8, and MAC1) that were established from G418 resistant cell colonies and analyzed by Northern blot analysis, which confirmed replication of HCV (Figure 4).

Cell-derived HCV RNA is more efficient than in vitrotranscribed RNA in initiating replication in HeLa and mouse hepatoma cells.

Results showed that replication of HCV subgenomes in HeLa and mouse cells led to the selection of replicons with several novel mutations. The majority of these mutations were located in the NS3, NS4B, and NS5A regions. Moreover, the results showed that cell-derived RNA carrying some or all of these mutations was much more efficient in establishing G418-resistant colonies in HeLa cells than was RNA derived from Huh7 cells (Table 1).

10

Based on these observations, it was surmised that introduction of these mutations into available subgenomic replicons should alter or expand their tissue and host tropism. To test this hypothesis, 13 subgenomic replicons 15 were designed that carried mutations in NS3, NS4B, and NS5A alone or in combination with each other as described in Table 3. Of the 13 constructs examined, only two, pZS2 and pZS25, yielded a small number of G418- resistant colonies in HeLa cells (Table 4). Viral RNA replication 20 was confirmed by Northern blot analysis of total RNA isolated from six cell lines derived from those colonies. None of the variants yielded colonies in Hepal-6 cells. Moreover, negative-control experiments with in vitrotranscribed RNA derived from a variant containing a 25 frameshift mutation in NS5B did not yield any colonies that could be expanded into cell lines. Notably, save for one, all replicons were permissive for replication in Huh7 cells, albeit with significantly different efficiencies (Table 3). Interestingly, both pZS2 and - 30 pZS25 carried mutations in NS4B that were conserved in replicons from two independent HeLa cell lines, SL1 and SL2. In addition, these replicons had the S2204I mutation in NS5A that was previously found to be one of the most potent adaptive mutations for HCV replication in Huh7 35

cells. Because both replicons replicated very efficiently in Huh7 cells, the results suggested that the NS4B mutations could have contributed to the observed expansion of the tissue tropism of HCV replicons. In support of this hypothesis, the subgenome with the highest efficiency in Huh7 cells, pZS11 lacking mutations in NS4B (Table 3), did not yield any colonies in HeLa cells.

10 TABLE 3. Colony formation efficiency of mutant replicons in Huh7 cells

Vector <sup>i</sup>	Conserved mutation(s)	Mean (SD) CFE/ µg of RNA <sup>g</sup>	NS protein(s) affected
I <sub>377</sub> /NS3-3'	NA <sup>h</sup> .	2.3 (1.5)	NA
pZS10	A1ª	3.3 (3.5)	NS3
pZS1	C1 <sup>b</sup>	$1.4 \times 10^3 (3.9 \times 10^2)$	NS5A
pZS20	$\mathtt{B}^{\mathtt{c}}$	1.7 (0.6)	NS4AB
pZS11	A1 + C1	$2.4 \times 10^5 \ (7 \times 10^4)$	NS3, NS5A
pZS2	B + C1	$7.8 \times 10^4 \ (1.7 \times 10^4)$	NS4AB, NS5A
pZS12	A1 + B	0.3 (0.6)	NS3, NS4AB
pZS5	A1 + B + C1	15 ·	NS3, NS4AB, NS
pZS4	$B + C1 + A2^d$	165 (21)	NS3, NS4AB, NS
pZS8	A2 + B	0	NS3, NS4AB
pZS6	$A2 + B + C2^{e}$	8 (4)	NS3, NS4AB, NS
pZS15	C3 <sup>f</sup>	491 (183)	NS5A
pZS25	B + C3	$1.7 \times 10^4 (2.6 \times 10^3)$	NS4AB, NS5A
pZS45	A2 + B + C3	19 (2)	NS3, NS4AB, NS

a A1, mutation E1202G.

<sup>&</sup>lt;sup>b</sup> C1, mutations S2204I and D2254E.

<sup>&</sup>lt;sup>c</sup> B, mutations L1701F, Q1720R, Q1727H, and V1749A.

d A2, mutations E1202G, S1128A, and S1323P.

<sup>35 °</sup> C2, mutations S2204I, D2254E, R2290I, and I2324R.

 $<sup>^{\</sup>rm f}$  C3, all the mutations of C2 plus the deletion 2371-2413.

<sup>&</sup>lt;sup>g</sup> CFE, colony formation efficiency. Values are derived from three independent transfections of each replicon RNA.

h NA, not applicable.

<sup>40 \*</sup> Sequence ID Numbers for subgenomic replicons are listed in Table 6

TABLE 4. Colony formation efficiency of in vitrotranscribed RNA<sup>a</sup>

	of cDNA y (cell line smid)	No. of col Huh7	onies in tra HeLa	ansfected cells Hepal-6
GS4.1 10 SL1 MH4 pZS2 pZS25		>10,000 >10,000 >10,000 >10,000 >10,000	0, 0, 0 0, 3, 2 3, 4, 0 2, 3, 0 0, 2, 1	0, 1, 0 0, 1, 0 17, 0, 0 0, 0, 0 0, 0, 0

a Results from transfection experiments with in vitro-transcribed RNA from pooled clones isolated from the indicated cell line and from in vitro-transcribed RNA from pZS2 and pZS25 (Table 3).

To further explore the basis for the observed low 20 colony formation efficiency of in vitro-transcribed RNA in HeLa cells, it was determined if replication in HeLa cells led to the selection of adaptive mutations that were not discovered previously when cDNA clones from SL1 and SL2 cells were sequenced. For this purpose, cDNA 25 clones were isolated from total RNA obtained with pZS2and pZS25-derived cell lines, respectively. Nucleotide sequence analysis of both cDNA clones did not reveal any additional consensus mutations, suggesting that the two subgenomes were sufficiently adapted for replication in 30 HeLa cells (results not shown). However, as mentioned above, it was possible that a minor population of subgenomic replicons with additional mutations were present in these cell lines. To overcome this problem, a method for the isolation and cloning of cDNAs spanning 35 the NS3 to NS5B region was developed (see Materials and Methods). Replicon cDNA libraries were produced from GS4.1, SL1, and MH4 cells. Approximately 2,000 cDNA

clones were pooled and subsequently used for in vitro transcription of subgenomic RNA. With Huh7 cells, the colony formation efficiency of the pooled clones was comparable to that of the most efficient subgenomes, such as pZS2 or pZS25, and did not vary significantly with the 5 origin of the total RNA used for cDNA cloning (Table 4). Consistent with previous results, colony formation in HeLa and mouse cells was origin dependent, i.e., save for one case, colonies were observed only with clones derived from SL1 and MH4 cell lines. Notably, with this strategy 10 G418-resistant colonies were obtained for the first time with Hepal-6 cells by using in vitro-transcribed RNA. To confirm the presence of viral RNA, 11 colonies were expanded and Northern blot analysis was performed with total RNA. All 11 RNA samples analyzed contained viral 15 RNA ranging from approximately 0.1 to 1 ng/5  $\mu g$  of total RNA (results not shown).

Taken together, the results supported the hypothesis that mutations identified in subgenomic replicons expressed in HeLa and mouse cells play a role in adaptation of the replicons to certain cell-type-specific conditions.

#### DISCUSSION

20

25 HCV is known as a species- and tissue-specific virus. The results described herein show that replication of HCV can occur in cells derived from tissues other than liver, indicating that cellular factors required for RNA replication are expressed in cell types other than hepatocytes. One interpretation of this result is that the apparent tropism of HCV for hepatocytes is determined primarily at the level of virus entry or assembly or, alternatively, that HCV can infect many other tissues but has escaped detection due to very low amounts of RNA replication or accumulation. Extrahepatic tissues could

serve as reservoirs for HCV that, as with human immunodeficiency virus, could provide a source of viruses that are refractory to antiviral therapy and, importantly, can be responsible for infection of liver grafts following orthotopic liver transplantation (Chun, T. W., et al., 2002. J. Infect. Dis. 185:1672-1676; Laskus, T., et al., 2002. J. Infect. Dis. 185:417-421). Such a scenario would have profound implications for antiviral therapy. For example, the targeting of drugs to secondary sites of viral replication and the analysis of drug metabolism in cells other than hepatocytes would become important factors for the development of successful antiviral therapies.

It is conceivable that HCV quasispecies in hepatocytes and other tissues exhibit differences in 15 their composition due to the selection of variants with cell-type-specific adaptations. As shown in this Example, replication of subgenomes in HeLa cells led to the accumulation of clusters of mutations in the NS3, NS4B, and NS5A regions including a deletion in NS5A (Figure 2). 20 Mutations and deletions in NS5A have been found previously in genomes that replicated in Huh7 cells, which could suggest that expression of the natural form of this protein in cell culture somehow interferes with RNA replication (Blight, K. J., et al. 2000. Science 25 290:1972-1975; Guo, J. T., et al., 2001., J. Virol. 75:8516-8523; Ikeda, M., et al., 2002. J. Virol. 76:2997-3006; Lohmann, V., et al., 2003. J. Virol. 77:3007-3019; Lohmann, V., et al., 2001. J. Virol. 75:1437-1449). However, mutations in the amino terminus of NS4B have 30 previously not been observed. Notably, in both SL1 and SL2 cells, the mutations changed two or one glutamine residues, respectively, to one of the two basic amino acids arginine and histidine. Moreover, the mutation V1749A was present in all five cell lines examined (Table 35

2 and Figure 2). Thus far, these results show that these mutations appear to be required for replication in HeLa cells, because only replicons pZS2 and pZS25 carrying these mutations yielded colonies after transfection with in vitro-transcribed RNA (Tables 3 and 4). The amino terminus of NS4B is predicted to reside on the cytoplasmic side of endoplasmic reticulum membranes and may interact with other host or viral proteins required for RNA replication (Hugle, T., et al., 2001. Virology 284:70-81). As an integral endoplasmic reticulum membrane 10 protein, NS4B might provide a scaffold for the assembly of replication complexes and act as a regulator for RNA replication. More importantly, a recent study revealed that NS4B can induce particular membrane structures, called membranous webs, proposed to be the site for HCV 15 replication (Egger, D., et al., 2002. J. Virol. 76:5974-5984). Interestingly, genetic analyses with an HCVrelated pestivirus identified the amino-terminal region of NS4B as a determinant for cytotoxicity caused by high levels of virus replication (Qu, L., et al., 2001. J. 20 Virol. 75:10651-10662).

In summary, this example demonstrates that HCV RNA replication is not restricted to the human hepatoma cell line Huh7 but instead occurs in HeLa cells and hepatoma cells derived from mice. These findings further facilitate development of a mouse model for HCV infection.

# Example II RNA Polymerase Inhibitors Inhibit HCV 30 replication in Transformed HeLa Cells

25

35

In this example, the anti-HCV activity of 2'-C-methyladenosine (2CMA, an HCV RNA polymerase inhibitor) was tested on GS4.1 (Huh7) cells, and SL1 (HeLa). The cells were treated with 10  $\mu$ M 2CMA. The cells were

5

20

25

30

35

harvested at 6, 12, 24, 48, and 72 hours. Total cellular RNA was extracted and viral RNA (vRNA) analyzed by Northern blot analysis. These results indicate that 2CMA effectively inhibits HCV in GS4.1 and SL1 cells (Figure 5).

Next, the antiviral activity of the HCV RNA polymerase inhibitor 5-OH-cytidine was tested. GS4.1 (Huh7) cells and SL1 (HeLa cells were treated with the indicated amounts of 5-OH-cytidine. The DNA polymerase inhibitor 5-OH-deoxy-cytidine was used as a negative control. The cells were harvested 72 hours after incubation with the drugs. Total cellular RNA was extracted and viral RNA analyzed by Northern blot analysis. The intensity of the bands corresponding to HCV RNA was determined with a Fuji phosphoimager. These results indicate that 5-OH-cytidine effectively inhibits HCV replication in GS4.1 and SL1 cells (Figure 6).

## Example III Cytopathic and NonCytoPathic Responses in Cells Expressing Hepatitis C Virus

Currently, combination treatment with alphainterferon and ribavirin is the therapy of choice for HCV infection. But this treatment is not always effective, and other treatment choices are limited, or have unproven efficacy. Study of the mechanism of action of IFN- $\alpha$  may help elucidate a new, effective treatment for HCV, or help determine what makes HCV treatment effective.

DNA microarray studies revealed that the antiviral response induced by IFNs alters the expression of hundreds of genes and, hence, is far more complex than previously anticipated (Der, S. D., et al., 1998. Proc. Natl. Acad. Sci. USA 95:15623-15628). Little is known about the nature of the cellular proteins that specifically target viral components and, hence, are responsible for the inhibition of viral replication in

the absence of cell death. In contrast, the major signal transduction pathways required for the innate immune response against many viruses have been elucidated. The first wave of IFN-induced genes depends on the phosphorylation of STAT1 and STAT2 and their interaction with IRF9 (p48) to form the transcription factor complex In addition, viral double stranded RNA (dsRNA) and other unknown viral factors are believed to play an important role in the establishment of an antiviral state. They can activate dsRNA-dependent enzymes such as 10 protein kinase R (PKR) and 2',5'-oligoadenylate synthase (OAS), as well as other still-elusive protein kinases (Smith, E. J., et al., 2001. J. Biol. Chem. 276:8951-8957). IFN- $\alpha$  can induce a noncytopathic antiviral response or, alternatively, trigger apoptotic programs 15 leading to the elimination of infected cells (Tanaka, N., et al., 1998. Genes Cells 3:29-37).

Nucleotide sequence analyses of HCV genomes isolated from Japanese patients revealed a correlation between the presence of mutations in a short segment of NS5A, termed 20 the IFN sensitivity-determining region (ISDR), and resistance to antiviral therapy with IFN- $\alpha$  (Enomoto, N., et al., 1995. J. Clin. Investig. 96:224-230; Enomoto, N., et al., 1996. N. Engl. J. Med. 334:77-81). Subsequently, it was reported that the ISDR motif can bind to PKR 25 (Gale, M. J., Jr., et al., 1997. Virology 230:217-227). Importantly, the ISDR from IFN-resistant, but not from IFN-sensitive, HCV isolates appeared to be a substrate for PKR, suggesting that IFN treatment of chronically infected patients can lead to the selection of HCV 30 variants with ISDRs that can bind and inactivate PKR (Gale, M. J., Jr., et al., 1998. Clin. Diagn. Virol. 10:157-162; Tan, S. L., and M. G. Katze. 2001. Virology 284:1-12).

35 Accordingly, study of HCV variants and the pathway

by which IFN inhibits HCV is necessary to provide new HCV treatments, and to prevent selection of IFN- $\alpha$  resistant variants.

#### 5 MATERIALS AND METHODS

10

Chemicals and reagents. Recombinant IFN- $\alpha$ 2b (intron A) was purchased from Schering-Plough. Cycloheximide, 2-aminopurine (2-AP), genistein, sodium salicylate, and wortmannin were obtained from Sigma. SB 203580, PD 98059, vanadate, PP2, rapamycin, and lactacystin were obtained from Calbiochem. Epoxomicin was obtained from Boston Biochem, and caspase inhibitor ZVAD- fluoromethyl ketone (ZVAD-FMK) was obtained from Enzyme Systems Products.

- Cell culture. Huh7 and HeLa cells were maintained in 15 Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, penicillin G, streptomycin, nonessential amino acids, and L-glutamine. For the cell lines carrying HCV and Kunjin virus replicons, 500 µg of G418/ml was added to the medium. The GS4.1 cell line was -20 derived from a subclone of FCA1 cells as described previously (Guo, J. T., et al., 2001. J. Virol. 75:8516-8523). SL1 is a HeLa cell line expressing HCV subgenomic replicon I<sub>377</sub>NS3-3' (Lohmann, V., et al., 1999. Science 285:110-113; Zhu, Q., et al., 2003. J. Virol. 77:9204-25 9210). The KUNCD20 cells represent a pool of approximately 200 colonies of G418-resistant HeLa cells obtained after transfection with the Kunjin virus replicon C20DXrepNeo RNA (Khromykh, A. A., and E. G. Westaway. 1997. J. Virol. 71:1497-1505) (kindly provided 30 by A. Khromykh, Sir Albert Sakzewski Virus Research Center, Brisbane, Australia).
- Plasmids. pCMV-E3L expressing the vaccinia virus E3L protein was obtained from Robert Schneider, New York

University, New York. pln035 expressing virus-associated (VA) RNA was obtained from David Lazinski, Tufts University, Boston, Mass. pEF-HA-HPIV2 expressing the V protein of human parainfluenzavirus 2 (HPIV2) was obtained from Curt Horvath, Mount Sinai School of Medicine, New York, N.Y. To obtain cDNA clones of the gene encoding human Mx-1, Huh7 cells were treated with 100 IU of IFN- $\alpha/ml$  for 6 h and total cellular RNA was extracted with TRIzol reagent (Invitrogen) and firststrand cDNA was made with an oligo(dT)<sub>12-18</sub> primer and 10 Superscript II DNA polymerase (Invitrogen) by following the manufacturer's direction. For the amplification of Mx-A cDNA, the primers used were 5'-AGTATCGTGGTAGAGAGCTGC-3' (SEQ ID NO:15) and 5'-TAATACGACTCACTATAGGGATGTGGCTGGAGATGC-3' (SEQ ID NO:16). 15 The purified PCR fragment was cloned into the pGEM-T Easy vector (Promega). The identity of the cloned fragment was verified by nucleotide sequence analysis.

20 RNA extraction and Northern blot hybridization. Total cellular RNA was extracted with TRIzol reagent (Invitrogen) by following the manufacturer's direction. Five micrograms of total RNA was fractionated on a 1% agarose gel containing 2.2 M formaldehyde and transferred onto nylon membranes. Membranes were hybridized with riboprobes specific for plus-stranded HCV replicon RNA and Mx-A and β-actin mRNA in the conditions described previously (Guo, J. T., et al., 2001. J. Virol. 75:8516-8523).

30

35

Detection of eIF-2 $\alpha$  phosphorylation by Western blotting. For Western blot analysis of eIF-2 $\alpha$  phosphorylation, cells were treated with 100 IU of IFN- $\alpha$ /ml for 12 h and then transfected with 2 µg of poly(I:C) per 60-mm-diameter plate by using Lipofectamine (Invitrogen). After

3 h of incubation, cells were lysed in high-salt radioimmunoprecipitation assay buffer (50 mM Tris-HCl [pH 8.0], 250 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% sodium dodecyl sulfate [SDS]). Proteins (40 µg) were separated on SDS-10% polyacrylamide gel and 5 electrophoretically transferred to nitrocellulose membranes. Membranes were incubated with 50% methanol, washed extensively with water, and blocked with 3% casein in TNET buffer (10 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM EDTA, 0.05% Tween 20). Membranes were incubated with 10 rabbit polyclonal antibodies against  $eIF-2\alpha$  (a gift from Robert Schneider, New York University) or phosphorylated eIF-2 $\alpha$  (eIF-2 $\alpha$ -P; Research Genetics, Inc.) diluted in blocking solution for 1 h and then washed extensively with TNET buffer. Membranes were then incubated with 15 horseradish peroxidase-conjugated goat anti-rabbit and immunoglobulin G (IgG) (Amersham), respectively. The bound IgG was detected with Super-Signal chemiluminescence reagents (Pierce).

20

35

RPA. For the analysis of IFN-β gene expression, cells were treated with 100 IU of IFN-α/ml for 12 h and then transfected with 2 μg of poly(I:C) per 60-mm plate by using Lipofectamine (Invitrogen). After 3 h of incubation, total cellular RNA was extracted with TRIzol reagent (Invitrogen), and IFN-β mRNA levels were determined by RNase protection assay (RPA) with the help of the RPAII kit purchased from Ambion. The probes complementary to IFN-β (GenBank accession no. M25460) and β-actin (GenBank accession no. BC013380) mRNAs spanned positions 272 to 650 and 1030 to 1250, respectively.

Annexin V-FITC staining. SL1 cells were plated on coverslips in six-well plates 16 h prior to treatment. Cells were then mock treated or treated with 100 IU of

IFN-α/ml in the absence or presence of 20 μM ZVAD-FMK.
Coverslips were then put on slides and incubated with 100
μl of staining solution containing annexin V-fluorescein
isothiocyanate (FITC) at room temperature for 10 to 15
min. After extensive washes with phosphate-buffered
saline, slides were examined with a Nikon fluorescence
microscope and photographed with a charge coupled device
camera.

Flow cytometry analysis. To determine the fraction of 10 apoptotic cells, the annexin V assay system (Roche Diagnostics GmbH) was used. Cells were incubated with IFN- $\alpha$  (100 IU/ml) alone or together with 20  $\mu$ M ZVAD-FMK for 24 h. The culture medium containing detached cells was collected, and the adherent cells were trypsinized 15 and then combined with the detached cells. The cells were collected by centrifugation and washed once in phosphatebuffered saline. Pelleted cells were resuspended in binding buffer and were incubated with annexin V-FITC at room temperature for 10 to 15 min. The stained cells were 20 then diluted with binding buffer and analyzed by flow cytometry (FACScan; Becton Dickinson).

#### RESULTS

30

35

25 IFN-α can induce noncytopathic and cytopathic antiviral responses in cells comprising HCV replicons.

As set forth above, stable HeLa cell lines were established that express HCV subgenomes with an efficiency similar to that of Huh7 cells (Zhu, Q., et al., 2003. J. Virol. 77:9204-9210). Examination of the IFN- $\alpha$  response in HeLa derived cell lines such as SL1 revealed a very similar dose dependent reduction of virus replication. The IC50 of IFN- $\alpha$  in HeLa cell lines was generally in the range of 0.1 IU/ml, approximately 10-fold lower than the IFN- $\alpha$  IC50 in Huh7-derived cell lines

(Figure 7). However, in marked contrast to observations made with Huh7 cell lines, treatment of SL1 and other HeLa derived cell lines with more than 30 IU of IFN- $\alpha/ml$ induced cell death in a significant fraction of cells between 6 and 20 h post treatment (Figure 8A). Cell death 5 was caused by apoptosis, as determined by annexin V staining, and could be prevented by the caspase inhibitor ZVAD-FMK. The fraction of apoptotic cells was determined before and after treatment with the cytokine. The results showed that IFN- $\alpha$  induced apoptosis in more than 30% of 10 SL1 cells compared with 6 to 7% in untreated SL1 and parental HeLa cells (Figure 8B). Several other HeLaderived cell lines were examined to assure that the results obtained with SL1 cells reflected a general property of HeLa cells expressing HCV subgenomes. 15 Moreover, to test more directly whether IFN-induced apoptosis was caused by viral replication, two methods were used to inhibit RNA replication in SL1 cells. The first was based on the observation that replication of HCV subgenomes is temperature sensitive and is inhibited 20 at 39°C (J. A. Sohn and C. Seeger, unpublished observations). The second method relied on the availability of an inhibitor of the viral RNA polymerase. Consistent with a role for viral replication in the induction of apoptosis, cell death could be prevented 25 when viral replication was inhibited by either incubation of the cells for 60 h at the elevated temperature or treatment with a viral polymerase inhibitor (Figure 8B and result not shown).

30 These results raised the question of whether IFNinduced apoptosis reflected a general property of HeLa
cells expressing viral replicons. To address this
problem, the IFN response against HCV in SL1 cells was
compared with that against the flavivirus Kunjin virus in
35 HeLa cells. For this purpose a pool of HeLa cells,

KUNCD20, expressing Kunjin virus subgenomic replicons
lacking the structural genes, similar to the HCV
subgenomic replicons was established (Khromykh, A. A.,
and E. G. Westaway. 1997. J. Virol. 71:1497-1505). The

Kunjin virus RNA levels in these cells were approximately
fivefold higher than those observed with HCV in SL1
cells. In contrast to results with SL1 cells, treatment
of KUNCD20 cells with different concentrations of IFN-α
only slightly inhibited viral replication (Figure 9).

10 Importantly, cell death in IFN-α- treated KUNCD20 cells was not detected either by light microscopy or annexin V staining (results not shown). These results indicated that IFN-induced apoptosis is a property of HCV-expressing HeLa cells rather than a general property of HeLa cells replicating viral RNA genomes.

In summary, these results demonstrated that IFN- $\alpha$  could induce noncytopathic as well as cytopathic antiviral programs in cells expressing HCV replicons in a concentration— and cell type-dependent fashion. Moreover, the results showed that this antiviral program was specific for HCV replicons. Importantly, the results suggested that HCV replication could induce an innate cellular response that, in combination with IFN- $\alpha$ , could lead to apoptosis.

25

30

35

20

## IFN- $\alpha$ inhibits HCV replication through the Jak-STAT signal transduction pathway

Information about the signal transduction pathways responsible for execution of the IFN response has generally been obtained with cells treated with high concentrations (100 to 1,000 IU/ml) of the cytokine and with fibroblasts and epithelial cells, most of which cannot, to date, support HCV replication. Moreover, a recent study by Schlaak, et al. (Schlaak, J. F., et al., 2002., J. Biol. Chem. 277:49428-49437) revealed that the

IFN response could vary in a cell type-dependent manner. In addition, it was found that slight changes in cell culture conditions had major effects on HCV replication. Therefore, the observation that replication of HCV in both Huh7 and HeLa cells could be inhibited with low concentrations of the cytokine warranted a more careful study of the pathways involved in the antiviral program against HCV.

5

To investigate the nature of the IFN-α response against the HCV replicon, drugs that were known to 10 inhibit specific components of selected signal transduction pathways that play a role in the antiviral response induced by IFN- $\alpha$  were used (Table 5). The current model for the signal transduction pathway induced by IFN- $\alpha$  predicts that the IFN receptor associated 15 tyrosine kinases Jak1 and Tyk2 are activated and, in turn, phosphorylate the transcription factors STAT1 and STAT2, which are required for the induction of the cellular antiviral program (Figure 10) (Sen, G. C. 2001., Annu. Rev. Microbiol. 55:255-281; Stark, G. R., et al., 20 1998. Annu. Rev. Biochem. 67:227-264). Incubation of GS4.1 cells with the tyrosine kinase inhibitor genistein suppressed the induction of the IFN-induced Mx-1 gene (Figure 11A and 11B). Similarly, genistein antagonized the IFN- $\alpha$  response against the HCV replicon. An increase 25 in the concentration of the drug from 100 to 300 µM led to a complete inhibition of the IFN response against HCV (results not shown). Consistent with this result, it was found that the V protein of HPIV2 blocked the IFN response. The V protein of HPIV2 induces the degradation 30 of STAT2 and, hence, inhibits the IFN-induced activation of gene expression (Parisien, J. P., et al., 2002. J. Virol. 76:4190-4198) (Figures 11C and 11D). IFN- $\alpha$ treatment of cells expressing HPIV2 led to a twofold reduction of viral RNA levels. When adjusted for the 35

observed transfection efficiency, i.e., 40 to 45% of the cells express the V protein, the reduction corresponded to a complete inhibition of the IFN- $\alpha$  response. Finally, the decline of viral RNA levels was reduced in GS4.1 cells with IFN- $\alpha$  and cycloheximide, indicating that de novo protein synthesis was required for an antiviral response against HCV replication (Table 5).

IFN- $\alpha$  can activate, in addition to the STAT pathway, MAPKs, including extracellular signal-regulated kinase, p38 MAPK, and phosphatidylinositol 3 (PI3)-kinase-Akt 10 pathways (David, M., et al., 1995. Science 269:1721-1723; Goh, K. C., et al., 1999. EMBO J. 18:5601-5608; Pfeffer, L. M., et al., 1997. Science 276:1418-1420). However, in contrast to genistein, SB 203580, sodium salicylate, and wortmannin, known inhibitors of p38 MAPK, NF-kB, and PI-3 15 kinase, respectively, did not inhibit the IFN response at detectable levels, suggesting that the two major ancillary signaling pathways activated by IFN-α were not directly involved in inhibiting HCV replication in Huh7 cells (Table 5; results not shown). 20

In summary, the results showed that inhibition of HCV replication with IFN- $\alpha$  depended on a functional Jak-STAT pathway (Figure 10). Hence, the results demonstrated that the IFN response against HCV was genuine and did not reflect an unspecific effect of the cytokine.

25

Does HCV replication induce an antiviral state in infected cells? A critical step in activation of innate immunity is the induction of an antiviral state by dsRNA or viral proteins (Figure 10) (Taniguchi, T., and A. Takaoka. 2002. Curr. Opin. Immunol. 14:111-116; tenOever, B. R., et al., 2002. J. Virol. 76:3659-3669). As shown above, evidence for such a virus-induced activation was also obtained from IFN-treated HeLa cells expressing HCV replicons. To investigate the nature of this HCV-induced

activation, the phosphorylation levels of eIF-2 $\alpha$  and expression of IFN- $\beta$  were determined. eIF-2 $\alpha$  is a substrate of PKR, which is activated by dsRNA that can accumulate as a consequence of viral RNA replication (Srivastava, S. P., et al., 1998. J. Biol. Chem. 273:2416-2423; Williams, B. R. 3 July 2001, posting date. Signal integration via PKR. Sci STKE 2001:RE2. [Online.]). IFN- $\beta$  gene transcription is activated through the coordinate actions of three families of transcription

5

- 10 factors NF-kB, IRF3, and ATF2, all of which are activated by dsRNA and/or certain viral proteins (Figure 10) (Peters, K. L., et al., 2002. Proc. Natl. Acad. Sci. USA 99:6322-6327; tenOever, B. R., et al., 2002. J. Virol. 76:3659-3669).
- 15 First the levels of eIF- $2\alpha$ -P in Huh7, GS4.1, HeLa, and SL1 cells was determined in the presence and absence of dsRNA and IFN- $\alpha$ . The results showed that eIF- $2\alpha$ -P levels were not significantly elevated in cells expressing HCV replicons (GS4.1 and SL1) compared with 20 those in their parental cells (Huh7 and HeLa) (Figure 12A). Similarly, incubation of cells with IFN- $\alpha$  did not induce eIF- $2\alpha$ -P levels. In contrast, transfection of cells with poly(I:C), mimicking dsRNA, augmented eIF-2 $\alpha$ -P levels, particularly in HeLa and SL1 cells. Similar results were obtained when cells were primed with IFN- $\alpha$ 25 prior to transfection with poly(I:C). These results were confirmed with several other cell lines derived from Huh7 and HeLa cells.

Second, the levels of IFN-β mRNA in the four cell
lines was determined under the same conditions described above. In agreement with the above results, viral replication alone was not sufficient to activate IFN-β gene expression in both Huh7- and HeLa-derived cell lines (Figure 12B). In Huh7 cells and GS4.1 and other Huh7-

35 derived cells expressing HCV replicons, only a weak

induction of IFN- $\beta$  was observed when cells were primed with IFN- $\alpha$  and then transfected with poly(I:C). In contrast, IFN- $\beta$  transcription was induced in HeLa and SL1 cells by poly(I:C) alone and particularly in combination with IFN- $\alpha$ . Remarkably, expression of IFN- $\beta$  could be induced by IFN- $\alpha$  alone in SL1 cells but not in HeLa cells. Similar results were obtained with the HeLaderived SL2 cell line (results not shown).

In summary, the results showed that, while both Huh7 and HeLa cells were competent to activate dsRNA-dependent signal transduction pathways, HCV replication alone was not sufficient to induce a detectable dsRNA response in these cells. This result could indicate that dsRNA either does not accumulate during HCV replication or cannot access PKR and other dsRNA binding proteins. Importantly, the results showed that, despite the apparent lack of biologically active dsRNA, viral replication could activate certain cellular signal transduction pathways that could cooperate with IFN-α to activate the transcription of the IFN-β gene.

The results presented above favored a model predicting that IFN- $\alpha$  inhibited HCV replication by a mechanism that was independent of dsRNA-activated antiviral pathways. To test this model more carefully, the IFN response was measured in GS4.1 cells expressing 25 the vaccinia virus E3L protein. E3L is known to sequester dsRNA and prevent PKR and OAS/Rnase L activation (Chang, H. W., et al., 1992. Proc. Natl. Acad. Sci. USA 89:4825-4829; Rivas, C., et al., 1998. Virology 243:406-414) (Figure 10). Indeed, expression of E3L had no measurable 30 effect on the IFN response against HCV (Figure 11C and 11D). Experiments relying on simultaneous detection of E3L and the HCV protein NS5A in the same cell by immunofluorescence confirmed that IFN- $\alpha$  could inhibit HCV replication in cells expressing the E3L protein (results 35

not shown). Finally, it was found that the PKR inhibitors 2-amino purine (2-AP) and adenovirus VA RNA did not block the IFN response against HCV in Huh7 cells (Table 5).

5 TABLE 5. Effects of inhibitors on the activity of IFN- $\alpha$  against the HCV replicon<sup>a</sup>

Drug, protein, or RNA	Concn	Primary target(s)	Effect
2-AP	10 mM	PKR and other kinases	
Genistein	300 µM	Tyrosine kinases	+
Cycloheximide	10 µg/ml	Translation	+
Sodium salicylate	5 mM	IKK	_
SB 203580	20 рм р38	MAPK	-
PD 98059	50 μM	MEK kinase	-
Vanadate	50 µM	Protein phosphatase	
Wortmannin	100 nM	PI3 kinase	_
PP2	50 μM	src kinase	-
Rapamycin	200 nM	mTOR, translation	•
Lactacystin	5 µM 26S	proteasome	+
Epoxomicin	1 µM 26S	proteasome	+
V protein of HPIV2		STAT2	+
E3L protein		PKR and OAS	-
VA RNA		PKR	-

a GS4.1 cells were incubated with the indicated compounds for 2 h and then the presence of 100 IU of IFN- $\alpha/ml$  for an additional 24 h. Viral RNA levels were determined by Northern blot analysis. Assays for V protein, E3L, and VA RNA are described in the legend to Fig. 5.

What are the pathways that play a role in the IFN-α response against HCV? A major question concerns the mechanism by which IFN-α induces the noncytopathic
35 inhibition of HCV replication. DNA microarray analyses of IFN-α-treated GS4.1 cells and other Huh7-derived cell lines revealed the induction of several classes of genes belonging to known signal transduction and protein degradation pathways (J. Hayashi and C. Seeger,
40 unpublished results; Cheney, I. W., et al. 2002. J. Virol. 76:11148-11154). In particular, several genes encoding proteasome subunits and ubiquitin-like proteins were among the genes most highly induced by IFN-α.

Notably, kinetic studies of HCV replication in Huh7 cells indicated that replication complexes have a relatively short halflife, which is further reduced by IFN treatment (J.-T. Guo and C. Seeger, unpublished observations).

Therefore, it was surmised that the proteasome could play a role in inhibition of HCV replication in IFN-treated cells.

To test this hypothesis, the outcome of combination treatment with IFN- $\alpha$  and the proteasome inhibitors lactacystin and epoxomicin for HCV RNA replication in 10 GS4.1 cells was determined. The cells were pretreated with different concentrations of the inhibitors for 1 h before IFN- $\alpha$  was added for an additional 6 h of combination treatment. Then the cells were incubated for 12 h before RNA was isolated and subjected to Northern 15 blot analysis (Figure 13A). The relatively short incubation period was necessary because of the known toxicity of proteasome inhibitors after longer incubation times. The results showed that HCV RNA levels dropped 70% within 18 h of IFN- $\alpha$  treatment, whereas in the presence 20 of epoxomicin or lactacystin the reduction was only 30% (Figure 13B). Lower doses of epoxomicin than of lactacystin were effective, which is consistent with the high specific activity of epoxomicin against the chymotrypsin-like activity of proteasomes (Fenteany, G., 25 and S. L. Schreiber. 1998. J. Biol. Chem. 273:8545-8548; Meng, L., et al., 1999. Proc. Natl. Acad. Sci. USA 96:10403-10408). Treatment with higher doses of lactacystin alone led to a slight reduction of HCV RNA levels. These results were confirmed with a second set of 30 experiments. The cells were pretreated with 5  $\mu M$ lactacystin and 1  $\mu M$  epoxomicin, respectively, for 1 h before IFN-q was added for an additional 6 h of combination treatment. RNA was isolated from the treated cells either 6 or 12 h after incubation with IFN- $\alpha$ 35

(Figure 14). The results showed that, at both time points, viral RNA levels were significantly higher in cells that were exposed to the proteasome inhibitors than in cells that were treated with IFN- $\alpha$  alone.

Finally, tests were conducted to determine whether proteasome activity was required for induction of the IFN response or, more directly, for inhibition of HCV replication. To distinguish between these two possibilities, first GS4.1 cells were incubated with the cytokine for 10 h to induce the antiviral response. Then, the cells were incubated for 12 h in the presence of lactacystin or epoxomicin. Under these conditions, the IFN response against HCV remained effective and reduced RNA levels to similar extents independently of the presence of either of the two inhibitors (Figure 15). Thus, these results indicated that the activity of proteasomes is required for the induction of the IFN response against HCV, but apparently not for direct inhibition of viral replication (see Discussion).

#### DISCUSSION

In this Example, the mechanism of the IFN- $\alpha$  response against subgenomic replicons of HCV in Huh7 and HeLa cells is investigated. The following conclusions can be drawn from these investigations. First, it can be concluded that IFN- $\alpha$  can inhibit HCV replication by both noncytopathic and cytopathic mechanisms. These results demonstrating that SL1 cells treated with IFN- $\alpha$  (100 IU/ml) underwent programmed cell death raised the question of whether apoptosis contributes to the rapid decline of HCV RNA levels observed during the first 48 h of IFN therapy (Neumann, A. U., et al., 1998. Science 282:103-107). The answer depends on whether HeLa or Huh7 cells mimic the scenario in HCV-infected hepatocytes in vivo. It is known from this and other studies that Huh7

cells exhibit an attenuated response to dsRNA and cannot induce an apoptotic program (results not shown) (Keskinen, P., et al., 1999. Virology 263:364-375; Lanford, R. E., et al., 2003. J. Virol. 77:1092-1104; McNair, A. N., et al., 1994. J. Gen. Virol. 75:1371-5 1378). In contrast, HeLa cells respond to dsRNA in a fashion similar to that in which primary chimpanzee hepatocytes respond. For example, treatment of chimpanzee primary hepatocyte cultures with poly(I:C) led to the induction of IFN- $\beta$ , as shown in this report with HeLa 10 cells (Figure 12) (Lanford, R. E., et al., 2003. J. Virol. 77:1092-1104). Therefore, it is likely that HeLa cells represent a more physiological model for hepatocytes in terms of IFN response than Huh7 cells. It was notable that only a fraction of SL1 cells died after 15 treatment with IFN- $\alpha$ . One possibility is that apoptosis was induced in cells that replicated above-average levels of HCV RNA. In support of this possibility, reduction of viral levels by treatment with heat or HCV RNA polymerase inhibitors reduced the number of apoptotic cells after 20 IFN-  $\alpha$  treatment (Figure 8 and results not shown). Based on these results it was concluded that HeLa cells did not undergo apoptosis by default after IFN- $\alpha$  treatment. In fact, it appears that apoptosis is a hallmark of HCVreplicating HeLa cells, because HeLa cells replicating 25 Kunjin virus RNA remained viable after IFN treatment. Finally, it appears that, the observation reported here represents the first example of IFN- $\alpha$ -induced apoptosis of cells replicating an apparently noncytolytic RNA 30 virus.

Second, it can be concluded that the noncytopathic response can occur independently of dsRNA-dependent pathways. Although these results showed that dsRNA response pathways were at least partially functional in normal and HCV-replicating Huh7 cells and were intact in

35

HeLa cells, as indicated by poly(I:C)-induced phosphorylation of eIF- $2\alpha$  and IFN- $\beta$  gene transcription, viral replication per se did not induce such responses (Figure 12). The expression of the vaccinia virus E3L protein or treatment of cells with the kinase inhibitor 2-AP had no measurable effect on the IFN response against the HCV replicon (Figure 11 and Table 5). These observations were consistent with the notion that dsRNAdependent antiviral pathways, such as PKR and RNase L pathways, were not involved in IFN-induced inhibition of 10 HCV replication in Huh7 cells. Whether they play a role in HeLa cells is not yet known. Efforts to express E3L in SL1 cells were not successful due to the apparent toxicity of the protein, and treatment of HeLa cells with 2-AP for more than 12 h induced apoptosis (results not 15 shown). Importantly, it is not yet known whether IFNinduced apoptosis in HCV-expressing cells is dependent on PKR or other dsRNA-dependent pathways (see below).

In summary, the results showed that, while both Huh7 and HeLa cells were competent to activate dsRNA-dependent signal transduction pathways, HCV replication alone was not sufficient to induce a detectable dsRNA response in these cells. This result could indicate that dsRNA either did not accumulate during HCV replication or was not accessible to PKR and other dsRNA binding proteins.

20

25

30

35

Third, it can be concluded that HCV replication can induce innate immune pathways. In SL1 and other HCV-expressing HeLa cell lines (results not shown), but not in normal HeLa cells, IFN- $\alpha$  induced the expression of IFN- $\beta$ . This indicates that HCV replication activated an unknown cellular factor, perhaps a viral activated kinase as proposed by Smith and colleagues (Smith, E. J., et al., 2001. J. Biol. Chem. 276:8951-8957), that, in turn, activated one or more transcription factors required for IFN- $\beta$  transcriptional activation. Candidates include

IRF3, NF-kB, and ATF-2 (Figure 10). Because normal HeLa cells did not undergo apoptosis after IFN- $\alpha$  treatment, it can be concluded that HCV expression directs the IFN- $\alpha$  response into a cytopathic process.

5

10

15

20

Fourth, it can be concluded that the antiviral activity of IFN-α against HCV depends, in part, on functional proteasomes. How IFN-induced antiviral programs inhibit viral replication noncytopathically is not yet understood. The results shown here demonstrate that, for HCV replication, proteasomes were required for this process. However, the idea that proteasomes were directly involved in inhibition of HCV replication by increasing the turnover of replication complexes or viral proteins was not supported by these results. Instead evidence was obtained that induction of the IFN response was dependent on degradation of one or several proteins. Previously, it has been shown (Li, X. L., and B. A. Hassel. 2001. Cytokine 14:247-252) that proteasome inhibitors attenuated the induction of certain IFNstimulated genes. Because epoxomicin and lactacystin did not inhibit induction of Mx-A (results not shown), which is dependent on activation of the Jak-STAT pathway for the formation of ISGF3, proteasomes might be involved in the induction of the second-wave IFN-stimulated genes. Such a model is consistent with results published

Such a model is consistent with results published previously by Li and Hassel (Li, X. L., and B. A. Hassel. 2001. Cytokine 14:247-252), who found that treatment of cells with proteasome inhibitors did not inhibit phosphorylation of STAT1 and binding of ISGF3 to DNA. As a consequence of these results, the number of IFN-induced genes that play a role in inhibition of HCV replication by IFN-α can be reduced to those that are repressed by epoxomicin.

An important implication of these results for 35 clinical IFN- $\alpha$  therapy and the pathogenesis of HCV

infections is that besides the noncytopathic antiviral effects, IFN- $\alpha$  might also induce apoptosis of HCVinfected hepatocytes. At first glance, this possibility might be discounted because drug-induced cell death could lead immediately to the destruction of the infected liver. However, it is possible that only a fraction of hepatocytes express levels of HCV high enough to activate an apoptotic program in the presence of the high levels of IFN- $\alpha$  that are used for antiviral therapy. In this scenario cell death would occur unnoticed. An important 10 consequence of such a scenario would be that cell killing could play a major role in the recovery from chronic HCV infections, similar to the situation in natural recovery from transient infections with woodchuck hepatitis virus, a model for hepatitis B virus infections (Guo, J. T., et 15 al., 2000. J. Virol. 74:1495-1505).

Table 6. Listing of Sequence ID Numbers

Sequence	Sequence ID Number
I <sub>377</sub> /NS3-3'	SEQ ID NO:1
pZS1	SEQ ID NO:2
pZS2	SEQ ID NO:3
pZS4	SEQ ID NO:4
pZS5	SEQ ID NO:5
pZS6	SEQ ID NO:6
pZ <b>S</b> 8	SEQ ID NO:7
pZS10	SEQ ID NO:8
pZS11	SEQ ID NO:9
pZS12	SEQ ID NO:10
pZS15	SEQ ID NO:11
pZS20	SEQ ID NO:12
pZS25	SEQ ID NO:13
pZS45	SEQ ID NO:14
Mx-A cDNA primer #1	SEQ ID NO:15
Mx-A cDNA primer #2	SEQ ID NO:16

While certain preferred embodiments of the present invention have been described and specifically exemplified above, it is not intended that the invention be limited to such embodiments. Various modifications may be made to the invention without departing from the scope and spirit thereof as set forth in the following claims.

#### WHAT IS CLAIMED IS:

10

25

1. A cell-line which replicates hepatitis C virus (HCV), wherein said cell line is selected from the group consisting of a non-human cell line and a human non-hepatic cell line.

- 2. The cell line of claim 1, wherein the human non-hepatic cell line comprises epithelial cells.
- 3. The cell line of claim 2, wherein the human epithelial cells are HeLa cells.
- 4. The cell line of claim 1, wherein the non-human cell line comprises mouse cells of hepatic origin.
  - 5. The cell line of claim 4, wherein the mouse cells are Hepa1-6 cells.
- 20 6. The cell line of claim 4, wherein the mouse cells are AML12 cells.
  - 7. A non-human, non-chimpanzee host animal comprising cells which replicate HCV.

8. The non-human host animal of claim 7, which is a mouse.

- 9. A method for producing a human non-hepatic cell 30 that replicates HCV, comprising:
  - a) obtaining total RNA from a human hepatic cell culture that replicates HCV, said total RNA comprising a selection marker which renders cells expressing said RNA resistant to a selection agent;
- 35 b) introducing the total RNA into human non-

hepatic cells; and

30

c) selecting those cells which grow in the presence of said selection agent and replicate HCV.

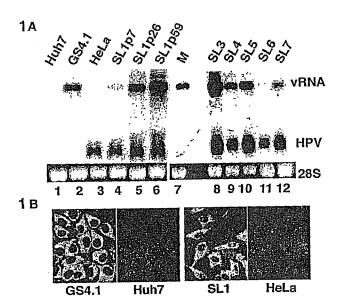
- 5 10. The method of claim 9, wherein a cell line is generated from the cells of step c).
  - 11. A method of producing a non-human hepatic cell that replicates HCV, comprising:
- a) obtaining total RNA from a human non-hepatic cell culture that replicates HCV, said total RNA comprising a selection marker which renders cells expressing said RNA resistant to a selection agent;
- b) introducing the total RNA into non-human 15 cells; and
  - c) selecting those cells which grow in the presence of said selection agent and replicate HCV.
- 12. The method of claim 11, wherein a cell line is 20 generated from the cells of step c).
  - 13. A method for screening test compounds which inhibit HCV replication, comprising:
- a) culturing the cell line of claim 1 in the 25 presence and absence of a test compound; and
  - b) assaying HCV replication levels in the presence and absence of said test compound, wherein a reduced HCV replication level in the presence of said test compound is indicative that said test compound inhibits HCV replication.
  - 14. An HCV polynucleotide having at least one of the mutations shown in Table II.
- 35 15. A polyprotein encoded by the polynucleotide of

claim 14.

- 16. A method for screening test compounds which modulate the antiviral response induced by interferon alpha (IFN- $\alpha$ ) comprising
- a) culturing the cell line of claim 1 in the presence and absence of a test compound;
  - b) contacting the cells of step a) with IFN- $\alpha$ ; and
- c) measuring the HCV replication level in the presence and absence of said compound thereby identifying agents which modulate the antiviral response mediated by IFN- $\alpha$  as a function of altered HCV levels.
- 17. The method of claim 16, wherein the antiviral response is enhanced.
  - 18. The method of claim 16, wherein the antiviral response is inhibited.

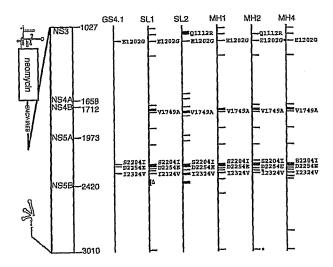
20

## Figures 1A-1B



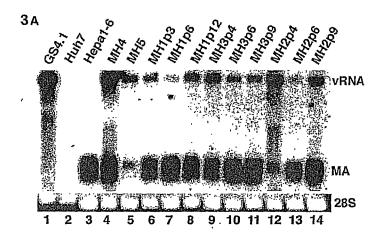
2/15

Figure 2

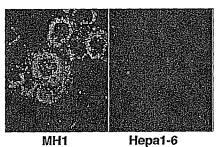


3/15

# Figures 3A-3B

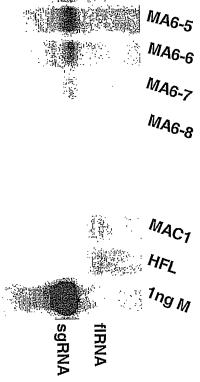


3 B

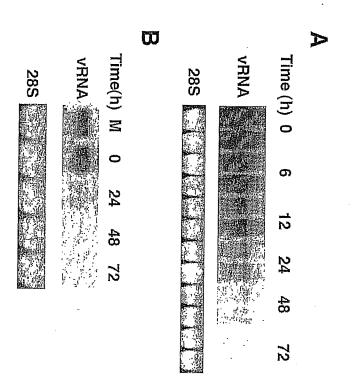


GS4.1

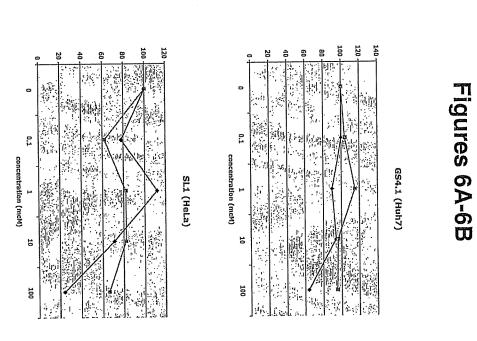






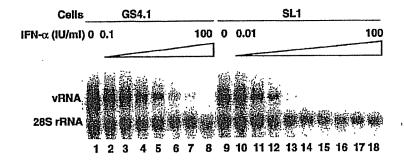


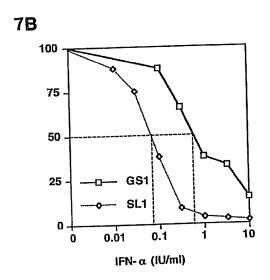
6**B** 



## Figures 7A-7B

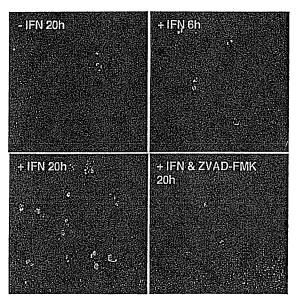
**7**A

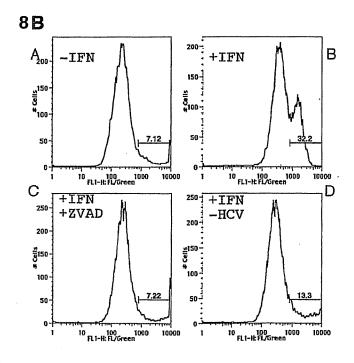




# Figures 8A-8B

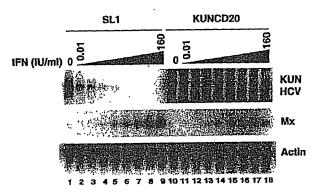
8A

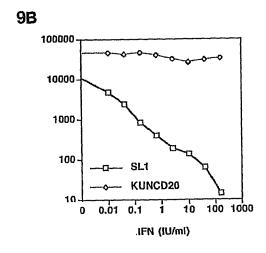




# Figures 9A-9B

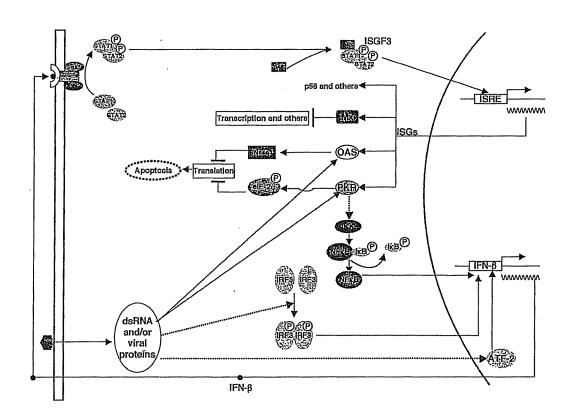
9A





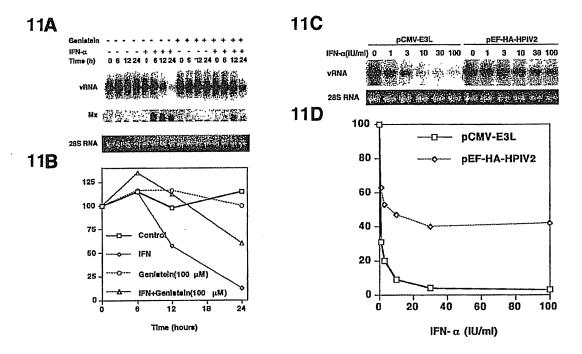
10/15

Figure 10

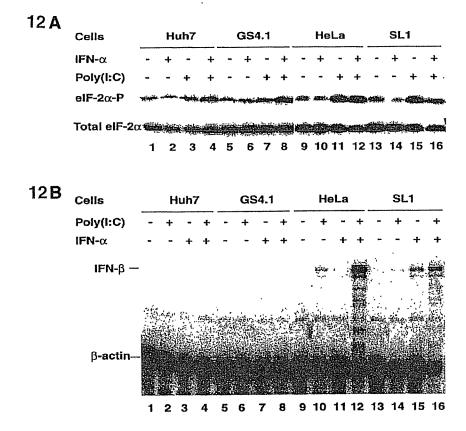


**3** 

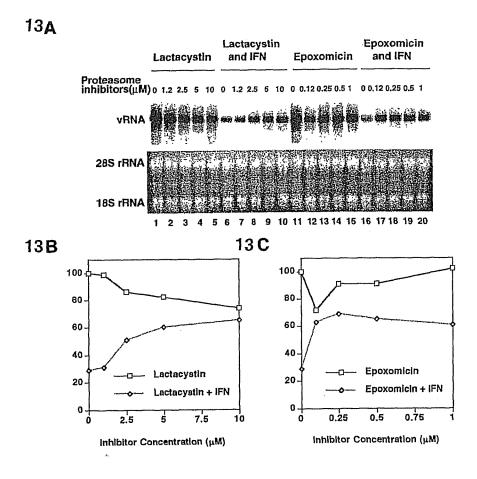
## Figures 11A-11D



## Figures 12A-12B

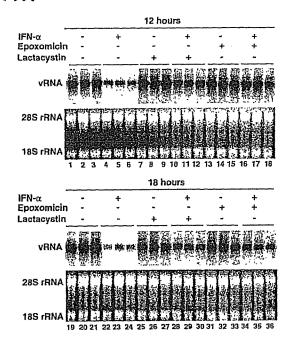


# Figures 13A-13C

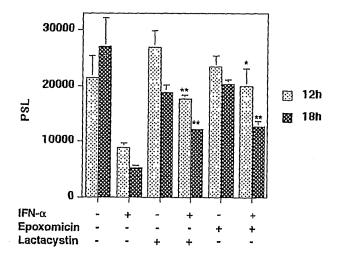


## Figures 14A-14B

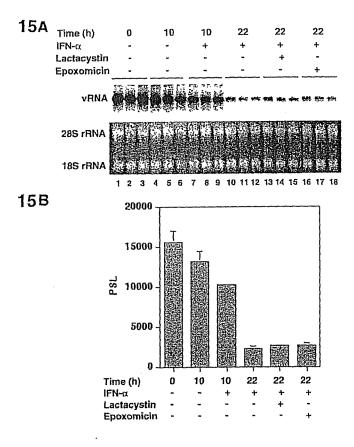
### 14 A



## 14B



## Figures 15A-15B



#### SEQUENCE LISTING

<110> Qing Zhu Ju-Tao Guo Christoph Seeger <120> Replication of Hepatitis C Virus in Non-Hepatic Epithelial and Mouse Hepatic Cells <130> 0149 P03068 <140> Not Yet Assigned <141> 2003-12-12 <150> 60/433,303 <151> 2002-12-13 <160> 16 <170> FastSEQ for Windows Version 3.0 <210> 1 <211> 11313 <212> DNA <213> Artificial Sequence <220>

<400> 1

<223> Plasmid

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 120 tettcacgca gaaagcgtet agccatggcg ttagtatgag tgtcgtgcag cetecaggae 180 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc 240 300 gegagaetge tageegagta gtgttgggte gegaaaggee ttgtggtaet geetgatagg 360 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 420 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 480 cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540 acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600 cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660 tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780 cattegacca ccaagegaaa categeateg agegageaeg taeteggatg gaageeggte 840 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900 ccaggeteaa ggegegeatg ceegaeggeg aggatetegt egtgaeeeat ggegatgeet 960 gettgeegaa tateatggtg gaaaatggee gettttetgg atteategae tgtggeegge 1020 1080 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140 agegeatege ettetatege ettettgaeg agttettetg agtttaaaea gaecacaaeg 1200 gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260 cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1320 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct 1380 aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440 gttcctctgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1500 aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560 gcaaaggcgg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa 1620 tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggt accccattgt 1680 atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gaggttaaaa 1740 aacgtctagg ccccccgaac cacggggacg tggttttcct ttgaaaaaca cgataatacc 1800 atggcgccta ttacggccta ctcccaacag acgcgaggcc tacttggctg catcatcact 1860 agcctcacag gccgggacag gaaccaggtc gagggggagg tccaagtggt ctccaccgca 1920 1980 acacaatett teetggegae etgegteaat ggegtgtgtt ggaetgteta teatggtgee 2040 ggctcaaaga cccttgccgg cccaaagggc ccaatcaccc aaatgtacac caatgtggac

caggacctcg	teggetggea	agcgcccccc	ggggcgcgtt	ccttgacacc	atgcacctgc	2100
aacaactcaa	acctttactt	ggtcacgagg	catgccgatg	tcattccggt	gcgccggcgg	2160
aacaacaaca	gggggagcct	actctcccc	aggcccgtct	cctacttgaa	gggctcttcg	2220
gacaatacac	tgctctgccc	ctcqqqqcac	actataaaca	tettteggge	tgccgtgtgc	2280
2222220000	ttgcgaaggc	gatagacttt	atacccatca	agtctatgga	aaccactatq	2340
caataccaa	tcttcacgga	caactcotco	cetecaacea	taccgcagac	attccaggtg	2400
cccccccgg	acgcccctac	taataacaac	aagaggacta	agataccaac	tacatataca	2460
accessace	ataaggtgct	tatactasac	ccatccatca	ccaccaccct	aggtttcggg	2520
gcccaagggc	ctaaggcaca	tootatooac	cctaacatca	gaaccggggt	aaggaccatc	2580
gegeacaege	ccccatcac	atactccacc	tatoocaadt	ttcttgccga	caataattac	2640
tataaaaaaa	cctatgacat	cataatatat	astasatacc	actcaactga	ctcgaccact	2700
ataataaaa	tcggcacagt	cataacacge	accaracac	ctagaacaca	actcatcata	2760
attectyggea	ctacgcctcc	agastagata	accataccac	atccasacat	casaasaata	2820
actatataa	gcactggaga	astagaettt	tataaaaaa	ccatccccat	cdadaccatc	2880
getetgteea	ggcacctcat	tttctccctt	tacaacaaa	astatastas	actracrara	2940
aagggggga	gcctcggact	gaatgatgata	cccaagaaga	aacgcgatga	totatocoto	3000
aagergreeg	gcggagacgt	caatgotgta	gcacaccacc	ctctaatcac	agactttacc	3060
acaccaacca	actcagtgat	agaatagaat	agatatata	ccccaacaat	caacttcacc	3120
ggcgatttcg	actiagigat	tgactgcaac	acatguguea	aagagagaga	atcacactca	3180
ecggaeeega	ccttcaccat	tagacyacy	ategreeattt	aagacgcggc	gecaegeeeg	3240
cageggegag	gcaggactgg	Laggggcagg	atgggcattt	acaggicigi	gaccccagga	3300
gaacggccct	cgggcatgtt	egatteeteg	greetteest	tagagaatta	agtagagaga	3360
gcttggtacg	agctcacgcc	egeegagace	ccagilaggi	rgegggerta	tagaggata	3420
ccagggttgc	ccgtctgcca	ggaccatctg	gagttetggg	agagegeeee	attaggeete	3480
acccacatag	acgcccattt	cttgtcccag	actaagcagg	caggagacaa	etteeectae	3540
ctggtagcat	accaggctac	ggtgtgcgcc	agggeteagg	ecedaectee	accycygyac	3600
caaatgtgga	agtgtctcat	acggctaaag	cetaegetge	acgggccaac	geceetgetg	
tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accccataac	caaacacacc	3660
atggcatgca	tgtcggctga	cctggaggtc	greacgagea	eetgggtget	ggtaggtgga	3720
gtcctagcag	ctctggccgc	gtattgcctg	acaacaggca	gegtggteat	tgtgggcagg	3780
atcatcttgt	ccggaaagcc	ggccatcatt	cccgacaggg	aagtccttta	ccgggagttc	3840
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacagggaat	gcagctcgcc	3900
gaacaattca	aacagaaggc	aatcgggttg	ctgcaaacag	ccaccaagca	agcggaggct	3960
	tggtggaatc					4020
tggaatttca	tcagcgggat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
gcgatagcat	cactgatggc	attcacagcc	tctatcacca	gcccgctcac	cacccaacat	4140
accctcctgt	ttaacatcct	ggggggatgg	grggccgccc	aacttgctcc	rcccagcgcr	4200
gcttctgctt	tcgtaggcgc	cggcatcgct	ggagcggctg	ttggcagcat	aggccttggg	4260
aaggtgcttg	tggatatttt	ggcaggttat	ggagcagggg	tggcaggcgc	geregragee	4320
tttaaggtca	tgagcggcga	gatgccctcc	accgaggacc	tggttaacct	actccctgct	4380
atcctctccc	ctggcgccct	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4440
gtgggcccag	gggagggggc	tgtgcagtgg	atgaaccggc	tgatagcgtt	cactrcacaa	4500
ggtaaccacg	tctcccccac	gcactatgtg	cctgagagcg	acgctgcagc	acgtgtcact	4560
cagatcctct	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gtggatcaac	4620
gaggactgct	ccacgccatg	ctccggctcg	tggctaagag	atgtttggga	ttggatatgc	4680
acggtgttga	ctgatttcaa	gacctggctc	cagtccaagc	teetgeegeg	attgccggga	4740
gtccccttct	tctcatgtca	acgtgggtac	aagggagtct	ggcgggggga	cggcatcatg	4800
caaaccacct	gcccatgtgg	agcacagatc	accggacatg	tgaaaaacgg	ttccatgagg	4860
atcgtggggc	ctaggacctg	tagtaacacg	tggcatggaa	cattccccat	taacgcgtac	4920
accacgggcc	cctgcacgcc	ctccccggcg	ccaaattatt	ctagggcgct	gtggcgggtg	4980
gctgctgagg	agtacgtgga	ggttacgcgg	gtgggggatt	tccactacgt	gacgggcatg	5040
accactgaca	acgtaaagtg	cccgtgtcag	gttccggccc	ccgaattctt	cacagaagtg	5100
gatggggtgc	ggttgcacag	gtacgctcca	gcgtgcaaac	ccctcctacg	ggaggaggtc	5160
	tcgggctcaa					5220
	cagtgctcac					5280
gctaagcgta	ggctggccag	gggatctccc	ccctccttgg	ccagctcatc	agctagccag	5340
ctgtctgcgc	cttccttgaa	ggcaacatgc	actacccgtc	atgactcccc	ggacgctgac	5400
	ccaacctcct					5460
tcagaaaata	aggtagtaat	tttggactct	ttcgagccgc	tccaagcgga	ggaggatgag	5520
	ccgttccggc					5580
	cacgcccgga					5640
	cagtggtaca					5700
	ggaagaggac					5760
	caaagacctt					5820
acggcctctc	ctgaccagcc	ctccgacgac	ggcgacgcgg	gatccgacgt	tgagtcgtac	5880
tcctccatgc	cccccttga	gggggagccg	ggggatcccq	atctcagcga	cgggtcttgg	5940
	gcgaggaggc					6000
_						

acaggcgccc	tgatcacgcc	atgcgctgcg	gaggaaacca	agctgcccat	caatgcactg	6060
agcaactctt	tactccatca	ccacaactta	gtctatgcta	caacaccccg	Caguguagu	6120
ctacaacaaa	agaaggtcac	ctttqacaqa	ctacaggice	Egyacyacca	ccaccyggac	6180 6240
atactcaago	agatgaaggc	gaaggcgtcc	acagttaagg	CLadactict	accegeggag	6300
daadeetata	agctgacgcc	cccacattcq	gccagaicia	aacceggeea	cggggcaaag	6360
gacgtccgga	acctatccag	caaggccgtt	aaccacatcc	gereegegeg	gatettetae	6420
ctggaagaca	ctgagacacc	aattgacacc	accaccacyy	tcotattccc	agatttgggg	6480
gtccaaccag	agaagggggg	ccgcaagcca ggccctttac	geregeerra	ccaccctccc	tcaggccgtg	6540
grrcgrgrgr	gegagaaaat	ccaatactct	cctggagaga	agatcaaatt	cctggtgaat	6600
argggererr	catacggatt	ccctatgggc	ttcgcatatg	acacccacta	ttttgactca	6660
accetaacta	agaagaaacg	ccgtgttgag	gagtcaatct	accaatgttg	tgacttggcc	6720
acggccaccg	gaacgacat	aaggtcgctc	acagagegge	tttacatcgg	gggccccctg	6780
actaattcta	aaggggagaa	ctacaactat	caccaatacc	gcgcgagcgg	tgtactgacg	6840
accadctdcd	gtaataccct	cacatgttac	ttgaaggccg	ctgcggcctg	Lugagulgug	6900
aagctccagg	actocacoat	actcatatac	ggagacgacc	ttgtcgttat	Cigigaaagc	6960
acadadaccc	aagaggacga	aacaaaccta	cgggccttca	cggaggctat	gactagatac	7020
tetacecec	ctaaaaaccc	gcccaaacca	gaatacgact	tggagttgat	aacaccacgc	7080
tectecaatq	tatcaatcac	gcacgatgca	tctggcaaaa	gggtgtacta	Lectacecyc	7140
gaccccacca	cccccttqc	acagactaca	tgggagacag	ctagacacac	Lecagicaac	7200 7260
tectaactaa	ocaacatcat	catgtatgcg	cccaccttgt	gggcaagyac	gateetgatg	7320
actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaaageeet	tanagacta	7380
atctacgggg	cctgttactc	cattgagcca	cttgacctac	ctcagatcat	taggacce	7440
catggcctta	gcgcattttc	actccatagt	tactetecay	gragarcaa	aaccaaaaat.	7500
tcatgcctca	ggaaacttgg	ggtaccgccc	regegageer	ggagacaccg	gtacctcttc	7560
gtccgcgcta	ggetaetgte	ccaggggggg gctcaaactc	agggccgcca	caactacata	ccaqttqqat	7620
aactgggcag	rattacttac	tggttacagc	addadadaca	tatatcacaq	cctatctcat	7680
gaaaaaaaaa	agtagttest	gtggtgccta	ctcctacttt	ctataggggt	aggcatctat	7740
geeegaeeee	accoatosac	ggggacctaa	acactccagg	ccaataggcc	atcctgtttt	7800
tttccctttt	ttttttttt	tttttttt	tttttttt	tttttttt	ttctcctttt	7860
tttttcctct	ttttttcctt	ttctttcctt	tggtggctcc	atcttagccc	tagtcacggc	7920
tacctctcaa	aggtccgtga	accacttaac	tqcaqagagt	gctgatactg	gcctctctgc	7980
agatcaagta	ctcctqcaqq	cgcgccacta	gtgggaatac	gcggggtatg	ccgcglllla	8040
gratattgac	gacccaattc	tcatqtttqa	cagettatea	tcgataagct	ctaatgeggt	8100
agtttatcac	agttaaattg	ctaacqcaqt	caggcaccgt	gtatgaaatc	taacaatgeg	8160 8220
ctcatcatca	tecteggeae	catcacccta	gatgctgtag	gcataggctt	ggreatgeeg	8220
gtactgccgg	gcctcttgcg	ggatatcgtc	cattccgaca	gcarcgccag	tcactatggc	8340
gtgctgctag	cgctatatgc	gttgatgcaa	tttctatgcg	tacttgaage	cggagcactg	8400
teegaceget	ttggccgccg	cccagtcctg	tagatagtat	accerggage	cactatcgac	8460
tacgcgatca	tggcgaccac	accegicacy	ggaccccc	tcaccaacat	catcgtggcc caccgatgg	8520
ggcatcaccg	gegeeacagg	caaactcata	agggettatt	tcaacataaa	caccgatggg tatggtggca	8580
gaagateggg	agagagaact	attacacaca	atctccttqc	atocaccatt	ccttgcggcg	8640
ggeeeegegg	accocctcaa	cctactacto	aactacttcc	taatqcaqqa	gtcgcataag	8700
geggegeeea	acggeeteaa	cttgagagcc	ttcaacccag	tcagctcctt	ccggtgggcg	8760
caaaacataa	ctategtege	cacacttata	actgtcttct	ttatcatgca	accegtagga	8820
cadataccad	cadedeteta	aatcattttc	ggcgaggacc	gctttcgctg	gagegegaeg	8880
atgatcggcc	tatcacttac	ggtattcgga	atcttgcacg	ccctcgctca	ageettegte	8940
actootccco	ccaccaaacq	tttcqqcqaq	aagcaggcca	ttategeegg	Carggeggee	9000
aacacactaa	getacqtett	actaacatto	qcqacqcqaq	gctggatggc	cttccccatt	9060
atgattcttc	taaattaaa	caacatcaaa	atgcccgcgt	tgcaggccat	getgteeayy	9120 9180
caggtagatg	acgaccatca	. qqqacaqctt	caaggatcgc	tegeggetet	Laccagecta	9180
acttcgatca	ctagaccact	gatcqtcacq	gcgatttatg	cegeeregge	gageacacgg	9300
aacgggttgg	catggattgt	aggcgccgcc	ctataccttg	tergeereee	cgcgttgcgt	9360
cgcggtgcat	ggagccgggc	cacctcgacc	rgaarggaag	ceggeggeac	ctcgctaacg	9420
gattcaccac	tccaagaatt	ggagccaacc	adcterigeg	gagaaccgcg	aatgcgcaaa	9480
ccaacccttg	geagaacaca	tecategege	tagacataat	catactecta	acgcggcgca tcgttgagga	9540
ccccgggcag	- ataaaaaaa	. tggccacggg	r attaccacaa	tgaatcacco	tcgttgagga	9600
geeggeragg	. cractoctoc	. tocasascot	ctacaeccta	agcaacaaca	tgaatggtct	9660
tecetttee	tatttate	agtotogaaa	cacaaaata	agcgccctgc	accattatgt	9720
tecagateta	catcacaga	ı tactactaac	: taccctgtgg	aacacctaca	. tctgtallaa	9780
cassacacta	r acattaacco	: taaataattt	: ttctctggtc	cegeegeare	: catacogoca	9840
attatttaca	: ctcacaacqt	: tccaqtaacc	: qqqcatqttc	e atcatcagta	acceguates	9900
tgagcatcct	ctctcgtttc	: atcggtatca	ttacccccat	. gaacagaaat	tececettae	9960
· · · · · · · · · · · · · · · · · · ·					and the second s	

acggaggcat	caagtgacca	aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
aagccagaca.	ttaacgcttc	tggagaaact	caacgagctg	gacgcggatg	aacaggcaga	10080
catctgtgaa	tcgcttcacg	accacgctga	tgagctttac	cgcagctgcc	tcgcgcgttt	10140
cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	cagcttgtct	10200
gtaagcggat	qccqggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	10260
tcggggggga	qccatqaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	10320
gcggcatcag	agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tqcqtaaqqa	gaaaataccg	catcaggcgc	tcttccgctt	cctcgctcac	tgactcgctg	10440
cgctcggtcg	ttcggctgcg	gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat	caggggataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaqqccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	10740
ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	10800
aggtatctca	gttcggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	10860
gttcagcccg	accqctqcqc	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	10920
cacqacttat	cqccactqqc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	10980
ggcggtgcta	cagagttctt	gaagtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttqqtatct	gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	11100
tccqqcaaac	aaaccaccgc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
cacaaaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	11220
tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
tagatccttt	tctagataat	acgactcact	ata			11313

<210> 2 <211> 11313 <212> DNA <213> Artificial Sequence <220> <223> Plasmid

<400> 2

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 tetteacgca gaaagegtet agceatggeg ttagtatgag tgtegtgeag cetecaggae 120 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180 gacgaccggg teetttettg gateaacccg etcaatgeet ggagatttgg gegtgeecce 240 300 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 360 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420 480 cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540 acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600 cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660 tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720 780 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg agcgagcacg tactcggatg gaagccggtc 840 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900 ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 960 gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1020 1080 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140 1200 agggcatcgc cttctatcgc cttcttgacg agttcttctg agtttaaaca gaccacaacg gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260 1320 cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1380 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct 1440 aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1500 gttcctctgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1560 aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1620 gcaaaggcgg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggt accccattgt 1680 atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gaggttaaaa 1740 aacgtctagg cccccgaac cacggggacg tggttttcct ttgaaaaaca cgataatacc 1800 atggcgccta ttacggccta ctcccaacag acgcgaggcc tacttggctg catcatcact 1860

agceteacag	accaggacag	gaaccaggtc	gagggggggg	tccaagtggt	ctccaccgca	1920
agagaatgtt	taataacaac	ctgcgtcaat	gacatatatt	ggactgtcta	tcatggtgcc	1980
acacaacccc	ccciggegae	cogogecaac	990909090	55atataaaa	anatataana	2040
ggctcaaaga	cccttgccgg	cccaaagggc	ccaatcaccc	adatytacac	caacgiggac	
caggacctcg	tcggctggca	agcgcccccc	qqggcgcgtt	ccttgacacc	atgcacctgc	2100
adasaataaa	acctttactt	ggtcacgagg	catoccoato	tcattccqqt	acaccaacaa	2160
ggcagcicgg	accectace.	ggccacgagg	cacgoogacg	aataattaaa	accepathece	2220
ggcgacagca	gggggagcct	actctccccc	aggeeegiei	CCLacityaa	gggctcttcg	
agagatacaa	tactctaccc	ctcggggcac	actataaaaa	tctttcgggc	tgccgtgtgc	2280
2202220000	++aaaaaaaa	ggtggacttt	atacccatca	agtctatgga	aaccactato	2340
accegagggg,	LLgcgaaggc	ggragacter	gracecarea	to a second		2400
cggtccccgg	tcttcacgga	caactcgtcc	cctccggccg	taccgcagac	arrecaygeg	
acceptate	accccctac	tggtagcggc	aagagcacta	aggtgccggc	tgcgtatgca	2460
9000000000	absorbant	tgtcctgaac	acatacataa	ccaccaccat	aggtttcggg	2520
geceaagggt	acaaygcycc	Lyccolyaac	cegeeegeeg	009000000	233222333	2580
gcgtatatgt	ctaaggcaca	tggtatcgac	cctaacatca	gaaccggggu	aaggaccatc	
accacaaata	cccccatcac	gtactccacc	tatqqcaagt	ttcttgccga	cggtggttgc	2640
**********	aatataaaat	cataatatgt	astasataca	actcaactga	ct.cgaccact.	2700
rergggggeg	Cocacyacac	cacaacacyc	gacgagegee			2760
atcctgggca	tcggcacagt	cctggaccaa	gcggagacgg	etggagegeg	actoglogic	
ctcaccacca	ctacqcctcc	gggatcggtc	accgtgccac	atccaaacat	cgaggaggtg	2820
~atatataa	adadtadada	aatccccttt	tatoocaaao	ccatccccat	cgagaggatg	2880
gerergreea	gcaccggaga	aaccccccc	cacggcaaag		aataaaaaa	2940
aaggggggga	ggcacctcat	tttctgccat	tccaagaaga	aatgtgatga	geregeegeg	
aagctgtccg	acctcaaact	caatgctgta	gcatattacc	ggggccttga	tgtatccgtc	3000
atagggggg	addasascat	cattgtcgta	acaacaaaca	ctctaatgac	gaactttacc	3060
alaccaacta	guggagaugu	carregera	9000099009		ggagttgagg	3120
ggcgatttcg	actcagtgat	cgactgcaat	acatgtgtca	Cocayacayı	cgaccccagc	
ctggacccga	ccttcaccat	tgagacgacg	accgtgccac	aagacgcggt	gtcacgctcg	3180
0.0000000000000000000000000000000000000	aasaasataa	taggggcagg	atogocattt	acaggtttgt	gactccagga	3240
caycyycyay	gcaggaccgg	caggggcagg		2043300030	acacacatat	3300
gaacggccct	cgggcatgtt	cgattcctcg	gttctgtgcg	agtgctatga	egegggeege	
acttaataca	ageteaegee	cgccgagacc	tcagttaggt	tgcgggctta	cctaaacaca	3360
ganagattag	acatatacas	ggaccatctg	gagttctggg	agagggtctt	tacaggcctc	3420
ccagggrege	cogeologica	ggaccacceg	2420000233	-3-3-3	attagaatea	3480
acccacatag	acgcccattt	cttgtcccag	actaagcagg	Caggagacaa	CLEGGGGGGAG	
ctggtagcat	accaggetae	ggtgtgcgcc	agggctcagg	ctccacctcc	atcgtgggac	3540
assatataas	antatataat	acggctaaag	cctacactac	acqqqccaac	acccctacta	3600
caaacycyga	agegeeeeae	acggeraaag		2033355	ganatagatg	3660
tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accedataac	Cadalacate	
atggcatgca	tatcaactaa	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	3720
ataataaaa	atataaaaaa	gtattgcctg	acaacaggca	acataatcat	tataaacaaa	3780
greerageag	CLCLGGCCGC	gracigorog		50505500	aggaaatta	3840
atcatcttgt	ccggaaagcc	ggccatcatt	cccgacaggg	aayteetta	ccgggagccc	
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacagggaat	gcagctcgcc	3900
ganganattan	2202022000	aatcgggttg	ctgcaaacag	ccaccaagca	agcagagact	3960
gaacaaccca	aacagaaggc	aaccyggccg		aattataaaa	annagatata	4020
getgeteeeg	tggtggaatc	caagtggcgg	acceregaag	correrggge	yaaycacacy	
tggaatttca	tcaqcqqqat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
adastsadst	cactuatuuc	attcacagcc	tctatcacca	acccactcac	cacccaacat	4140
gcgatagcat	Laceguegge		~~~~~~~~	sacttactac	taggegaget	4200
accetectgt	ttaacatcct	ggggggatgg	gradecace	aactigetee	Leccagegee	
acttctactt	tcqtaqqcqc	cggcatcgct	ggagcggctg	ttggcagcat	aggccttggg	4260
aaaataatta	tagatatttt	ggcaggttat	adaddaddad	taacaaacac	actcataacc	4320
aagguguuu	Lygacacccc	9904990040	3343443333	+aattaaaat	actocotoct	4380
tttaaggtca	tgagcggcga	gatgccctcc	accgaggacc	Lygitaacci	acticities	
atcctctccc	ctggcgccct	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4440
ataaaaaaaa	aaasaaaaa	tgtgcagtgg	atgaaccggc	tgatagcgtt	cacttcacaa	4500
gragacccaa	999499999	2202042022	aataaaaaa	addatadada	acatataact	4560
ggtaaccacg	tctcccccac	gcactatgtg	curyayayuy	acyclycayc	acgugucacu	4620
cagatcctct	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	grggarcaac	
gaggactgct	ccacaccata	ctccggctcg	taactaaaaq	atqtttqqqa	ttggatatgc	4680
54554445	atantttann	gaggagaga	cantecaage	tectaceaca	attgccggga	4740
acggtgttga	Cigalilicaa	gaddliggdlid	cagcccaage	coccycogog	~~~~	4800
gtccccttct	tctcatgtca	acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	
caaaccacct	acccatataa	agcacagatc	accggacatg	tgaaaaacgg	ttccatgagg	4860
	ataccadata	tagtaacacg	taacataaaa	cattegegat	taacqcqtac	4920
accycygygo	Claggacety	Lagitaacacg	cggcacggaa		apacacacta	4980
accacgggcc	cctgcacgcc	ctccccggcg	ccaaattatt	ctagggcgct	gtggcgggtg	
actactaaaa	agtacgtgga	gattacacaa	qtqqqqqatt	tccactacgt	gacgggcatg	5040
200200202	adatasaata	accatatasa	attecaacee	cccaattctt	cacagaagtg	5100
accaccyaca	acgraaagrg	cccgtgtcag	922239222		accadadta	5160
gatggggtgc	ggttgcacag	gtacgeteca	gcgtgcaaac	ecetectacy	ggaggaggtc	
acattcctoo	tcgggctcaa	tcaatacctq	gttgggtcac	agctcccatg	cgagcccgaa	5220
addas aat sa	cactoctosc	ttacetacta	accoacccc+	cccacattac	ggcggagacg	5280
coggacgtag	- uguguuda					5340
gctaagcgta	ggctggccag	gggatetece	eceteettgg	ccaycicatc	agctatccag	
ctatctacac	cttccttgaa	ggcaacatqc	actacccgtc	atgactcccc	ggacgctgac	5400
utustaasaa	acasacetact	ataaraaraa	asastanara	ggaadatdad	ccgcgtggag	5460
Licalogagg		2-22-23-42	5~5~55559	+	adaddataad	5520
tcagaaaata	aggtagtaat	tttggagtet	LLegageege	cccaagegga	ggaggatgag	
agggaagtat	ccgttccqqc	ggagatcctg	cggaggtcca	ggaaattccc	tcgagcgatg	5580
cccatatoro	daddddddda	ttacaaccct	ccactottag	agtectogaa	ggacccggac	5640
			++~~~~~+	-002244444	tocoatacca	5700
tacgtccctc	cagtggtaca	cgggtgtcca	Ligotycetg	Caayycccc	tccgatacca	
cctccacgga	ggaagaggac	ggttgtcctg	tcagaatcta	ccgtgtcttc	tgccttggcg	5760
gagetedeea	caaagacctt	caacaactcc	gaatcotcoo	ccqtcqacaq	cggcacggca	5820
5~500500		-550050000				

accordinate	ctgaccagcc	ctccgacgac	aacaacacaa	gatccgacgt	tgagtcgtac	5880
tectecatee	cccccttga	aaaaaaaacca	agaaatccca	atctcagcga	cagatettaa	5940
tatagataa	cccccccd	tagtgaggag	gtcatctact	gctcgatgtc	ctacacataa	6000
agragaga	tastascaca	atacactaca	daddaaacca	agctgcccat	caatgcactg	6060
acaggegeee	tgattacgee	acgogocycy	atctatacta	caacatctcg	cadcadaaad	6120
agcaaccccc	Lycicogcoa	atttaacaca	gtacaaataa	tggacgacca	ctaccoooac	6180
ctgcggcaga	agaaggicac	ctttgataga	cogcaggeee	ctasacttct	atccatacaa	6240
gtgctcaagg	agatgaagge	gaaggegeee	acageraagg	ctaaacttct	taaaaaaaa	6300
gaagcctgta	agetgaegee	cccacattcg	gecagateta	aatttggcta	cggggcaaag	6360
gacgtccgga	acctatccag	caaggccgtt	aaccacatcc	gctccgtgtg	gaaggacttg	
ctggaagaca	ctgagacacc	aattgacacc	accatcatgg	caaaaaatga	ggttttetge	6420
gtccaaccag	agaagggggg	ccgcaagcca	gctcgcctta	tcgtattccc	agatttgggg	6480
gttcgtgtgt	gcgagaaaat	ggccctttac	gatgtggtct	ccaccctccc	tcaggccgtg	6540
atgggctctt	catacqqatt	ccaatactct	cctggacagc	gggtcgagtt	cctggtgaat	6600
acctagaaag	cgaagaaatg	ccctatgggc	ttcgcatatg	acacccgctg	ttttgactca	6660
accotcacto	agaatgacat	ccgtgttgag	gagtcaatct	accaatgttg	tgacttggcc	6720
cccgaagcca	gacaggccat	aaggtcgctc	acagagcggc	tttacatcgg	gggccccctg	6780
actaattcta	aaqqqcaqaa	ctgcggctat	cgccggtgcc	gcgcgagcgg	tgtactgacg	6840
accadctdcd	gtaataccct	cacatgttac	ttgaaggccg	ctgcggcctg	tcgagctgcg	6900
aagctccagg	actocacoat	gctcgtatgc	ggagacgacc	ttgtcgttat	ctgtgaaagc	6960
acadadaccc	aagaggacga	aacaaaccta	cgggccttca	cggaggctat	gactagatac	7020
teteccece	ctagagaccc	gcccaaacca	gaatacgact	tggagttgat	aacatcatgc	7080
tectecaste	tatcaatcac	gcacgatgca	tctggcaaaa	gggtgtacta	tctcacccgt	7140
gagggagg	cccccttac	acadactaca	taggagagag	ctagacacac	tccagtcaat	7200
tactecatea	gangatast	catatataca	cccaccttat	gggcaaggat	gatectgatg	7260
	tetacateat	tatagataga	gaacaacttg	aaaaagccct	agattgtgag	7320
actcatttct	LCLCCalccl	anttangaga	attanantan	ctcacatcat	traacqactc	7380
acctacgggg	ectgttacte	cattgageca	toatataaaa	ctcagatcat	taggggggt	7440
catggcctta	gcgcattttc	actecatage	tacceccay	gtgagatcaa	cagggegget	7500
tcatgcctca	ggaaacttgg	ggtaccgccc	Lugegagiei	ggagacatcg	ggccagaagc	7560
gtccgcgcta	ggctactgtc	ccaggggggg	agggetgeea	cttgtggcaa	gracetteest	7620
aactgggcag	taaggaccaa	gctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	7620
ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	eetgtetegt	
gcccgacccc	gctggttcat	gtggtgccta	ctcctacttt	ctgtaggggt	aggeatetat	7740
ctactcccca	accgatgaac	ggggacctaa	acactccagg	ccaataggcc	atcctgttt	7800
tttccctttt	tttttttctt	tttttttt	tttttttt	tttttttt	ttctcctttt	7860
tttttcctct	ttttttcctt	ttctttcctt	tggtggctcc	atcttagccc	tagtcacggc	7920
tagctgtgaa	aggtccgtga	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	7980
agatcaagta	ctcctqcaqq	cqcqccacta	gtgggaatac	gcggggtatg	ccgcgtttta	8040
gcatattgac	gacccaattc	tcatgtttga	cagcttatca	tcgataagct	ttaatgcggt	8100
agtttatcac	agttaaattg	ctaacgcagt	caggcaccgt	gtatgaaatc	taacaatgcg	8160
ctcatcatca	tecteggeae	cgtcaccctg	gatgctgtag	gcataggctt	ggttatgccg	8220
atactaccaa	acctcttaca	ggatatcgtc	cattccgaca	gcatcgccag	tcactatggc	8280
atactactea	coctatatoc	attaatacaa	tttctatgcg	cacccgttct	cggagcactg	8340
tccaaccact	ttaaccacca	cccaatccta	ctcqcttcqc	tacttggagc	cactatcgac	8400
taccccatca	taacaaccac	acccatccta	tagatectet	acgccggacg	catcgtggcc	8460
addataaca	acaccacaa	tacaattact	ggcgcctata	tcaccaacat	caccgatggg	8520
ggcaccaccg	ctccccactt	caaactceta	agcgcttgtt	tcaacataaa	tatggtggca	8580
gaagaccggg	ccaaaaact	attagacaca	atctccttqc	atocaccatt	ccttgcggcg	8640
ggccccgcgg	accacacatas	cctactacta	gactacttcc	taatqcaqqa	gtcgcataag	8700
geggegeeea	acggccccaa	cttaecaecg	ttcaacccaq	tcagctcctt	ccggtgggcg	8760
ggagagegee	gaccgacgcc	accasattata	actatatat	ttatcatcca	actcgtagga	8820
cggggcacga	crarcyccyc	getacttate	accascasca	actttcacta	gagcgcgacg	8880
caggtgccgg	cagegeterg	ggtcatttt	ggcgaggacc	ccctcgctca	adocttoato	8940
atgateggee	tgregerige	ggtattegga	accetgeacg	ttatagaaaa	catoggggg	9000
actggtcccg	ccaccaaacg	ttteggegag	aagcaggcca	ttatcgccgg	cttccccatt	9060
gacgcgctgg	geraegrerr	getggegtte	gegaegegag	gctggatggc	actatacaaa	9120
atgattcttc	tegetteegg	cggcarcggg	atgeeegegt	tgeaggeeat	gctgtccagg	9180
caggtagatg	acgaccatca	gggacagctt	caaggatege	regeggerer	taccagccta	9240
acttcgatca	ctggaccgct	gatcgtcacg	gegatttatg	eegeeregge	gagcacatgg	9300
aacgggttgg	catggattgt	aggcgccgcc	ctataccttg	tctgcctccc	cycyttycyt	9360
cgcggtgcat	ggagccgggc	cacctcgacd	rgaarggaag	ccggcggcac	ctcgctaacg	
gattcaccac	tccaagaatt	ggagccaatc	aattcttgcg	gagaactgtg	aatgcgcaaa	9420
ccaacccttq	gcagaacata	tccatcgcgt	ccgccatctc	cagcagccgc	acgcggcgca	9480
tctcgggcag	cgttgggtcc	tggccacggg	tgcgcatgat	cgtgctcctg	tcgttgagga	9540
cccqqctagg	ctggcggggt	tgccttactg	gttagcagaa	tgaatcaccg	atacgcgagc	9600
gaacgtgaag	cqactgctgc	tgcaaaacgt	ctgcgacctg	agcaacaaca	tgaatggtct	9660
tcggtttccg	tgtttcgtaa	. agtctggaaa	. cgcggaagtc	agcgccctgc	accattatgt	9720
tccggatctg	categeagga	. tgctgctggd	taccctgtgg	aacacctaca	tctgtattaa	9780

cgaagcgctg	gcattgaccc	tgagtgattt	ttctctggtc	ccgccgcatc	cataccgcca	9840
gttgtttacc	ctcacaacgt	tccagtaacc	gggcatgttc	atcatcagta	acccgtatcg	9900
tgagcatcct	ctctcgtttc	atcggtatca	ttacccccat	gaacagaaat	tcccccttac	9960
acqqaqqcat	caagtgacca	aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
aagccagaca	ttaacqcttc	tggagaaact	caacgagctg	gacgcggatg	aacaggcaga	10080
catctqtqaa	tcqcttcacg	accacgctga	tgagctttac	cgcagctgcc	tcgcgcgttt	10140
cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	cagcttgtct	10200
gtaagcggat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	10260
tcggggggga	gccatgaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	10320
acaacatcaa	agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tacataaqqa	gaaaataccg	catcaggcgc	tcttccgctt	cctcgctcac	tgactcgctg	10440
cactcaatca	ttcggctgcg	gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat	caggggataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	cccctqqaaq	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	10740
ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	10800
aggtatctca	gttcggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	10860
attcaaccca	accactacac	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	10920
cacqacttat	cqccactggc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	10980
aacaatacta	cagagttctt	qaaqtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttggtatct	gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	11100
tccqqcaaac	aaaccaccqc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
cqcaqaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	11220
tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
		acgactcact				11313
_						

<210> 3 <211> 11313 <212> DNA <213> Artificial Sequence <220> <223> Plasmid

<400> 3

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 tottcacgca gaaagcgtot agccatggcg ttagtatgag tgtcgtgcag cotccaggac 120 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180 gacgaccggg teetttettg gateaacccg etcaatgeet ggagatttgg gegtgeecce 240 300 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420 cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480 540 ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600 cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660 tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780 cattegacea ecaagegaaa categeateg agegageaeg taeteggatg gaageeggte 840 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900 ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 960 1020 gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1080 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1140 ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1200 agcgcatcgc cttctatcgc cttcttgacg agttcttctg agtttaaaca gaccacaacg gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260 1320 cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1380 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440 1500 gttcctctgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1560 aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1620 qcaaaqqcqq cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggt accccattgt 1680

atoggateto	atctggggcc	tcqqtqcaca	toctttacat	gtgtttagtc	gaggttaaaa	1740
acgggacocg	ccccccaac	cacadaasca	taattttcct	ttgaaaaaca	cgataatacc	1800
atagaaata	ttacccccta	ctcccaacag	acacasaacc	tacttggctg	catcatcact	1860
acggegeeea	addadacaa	gaaccacctg	asaaaaaaaaa	tccaagtggt	ctccaccaca	1920
agecteacag	tactagacas	ctacatast	gagggggaga	ggactgtcta	tcatootgcc	1980
acacaacccc	Lectggegae	adassaccas	ggcgcgcgcc	asatrtarar	caatutuuac	2040
ggctcaaaga	cectigeegg	cccaaaggge	CCaaccaccc	aaatgtacac	atacacatac	2100
caggacctcg	teggetggea	agegeeece	ggggcgcgcc	ccttgacacc	acgeaececyc	2160
ggcagctcgg	acctttactt	ggccacgagg	catgeegacg	tcattccggt	gegeeggegg	
ggcgacagca	gggggagcct	actctcccc	aggcccgtct	cctacttgaa	gggctcttcg	2220
ggcggtccac	tgctctgccc	ctcggggcac	gctgtgggca	tctttcgggc	rgccgrgrgc	2280
acccgagggg	ttgcgaaggc	ggtggacttt	gtacccgtcg	agtctatgga	aaccactatg	2340
cggtccccgg	tcttcacgga	caactcgtcc	cctccggccg	taccgcagac	attccaggtg	2400
gcccatctac	acgcccctac	tggtagcggc	aagagcacta	aggtgccggc	tgcgtatgca	2460
gcccaagggt	ataaggtgct	tgtcctgaac	ccgtccgtcg	ccgccaccct	aggtttcggg	2520
acatatatat	ctaaqqcaca	tggtatcgac	cctaacatca	gaaccggggt	aaggaccatc	2580
accacqqqtq	ccccatcac	gtactccacc	tatggcaagt	ttcttgccga	cggtggttgc	2640
tctaaaaaca	cctatgacat	cataatatgt	gatgagtgcc	actcaactga	ctcgaccact	2700
atcctgggca	tcaacaat	cctqqaccaa	gcggagacgg	ctggagcgcg	actcgtcgtg	2760
ctcaccacca	ctacacctcc	aggatcagtc	accetecac	atccaaacat	cgaggaggtg	2820
actotatoca	gcactggaga	aatccccttt	tatqqcaaaq	ccatccccat	cgagaccatc	2880
acconductura	gacacctcat	tttctqccat	tccaagaaga	aatgtgatga	gctcgccgcg	2940
aaggggggaa	acctcaaact	caatoctota	gcatattacc	ggggccttga	tatatccatc	3000
atagazata	gccccggact	cattotcota	dcaacddacd	ctctaatgac	gaactttacc	3060
acaccaacca	geggagaege	agagtagaat	acatototoa	cccagacagt	cgacttcagc	3120
ggegattteg	acteagryat	tananaa	acatguguea	aadacacaat	atcacactca	3180
etggaeeega	ectteaceat	Lyayacyacy	accycyctat	aagacgcggt	geologoog	3240
cagcggcgag	gcaggactgg	taggggcagg	argggcarrr	acaggtttgt	gaccccagga	3300
gaacggccct	cgggcatgtt	cgarrectcg	gecetgegeg	agtgctatga	cgcgggccgc	3360
gcttggtacg	agctcacgcc	cgccgagacc	ccagttaggt	tgcgggctta	tocadacaca	3420
ccagggttgc	ccgtctgcca	ggaccatctg	gagttctggg	agagcgtctt	tacaggeete	
acccacatag	acgcccattt	cttgtcccag	actaagcagg	caggagacaa	CTTCCCCTac	3480
ctggtagcat	accaggctac	ggtgtgcgcc	agggctcagg	ctccacctcc	atcgtgggac	3540
caaatgtgga	agtgtctcat	acggctaaag	cctacgctgc	acgggccaac	gacactgatg	3600
tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accccataac	caaatacatc	3660
atggcatgca	tqtcagctga	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	3720
atcctagcag	ctctqqccqc	gtattgcctg	acaacaggca	gcgtggtcat	tgtgggcagg	3780
atcatcttqt	ccqqaaaqcc	ggccatcatt	cccgacaggg	aagtcttta	ccgggagttc	3840
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacggggaat	gcagctcgcc	3900
gaacatttca	aacagaaggc	aatcqqqttg	ctgcaaacag	ccaccaagca	agcggaggct	3960
actactacca	caataaaatc	caagtggcgg	accctcgaag	ccttctgggc	gaagcatatg	4020
tagaattta	tcagcgggat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
acastaacat	cactgatggc	attcacagec	tctatcacca	gcccgctcac	cacccaacat	4140
accetectat	ttaacatect	agaggatag	ataaccaccc	aacttgctcc	tcccagcgct	4200
acttatactt	tratagggg	cadcatcact	ggagggggtg	ttggcagcat	aggccttggg	4260
agatactta	tagetaget	aggaagttat	adadcaddd	tggcaggcgc	actcataacc	4320
aagguguug	tanagagaga	ggtaggttat	2242442222	tggttaacct	actccctact	4380
-t-caaggica	chagoggoga	actactaca	atcatataca	cagcgatact	acatcaacac	4440
atecteteee	erggegeeer	tatagaataa	atraaccooc	tgatagcgtt	cacttcacaa	4500
grgggcccag	gggagggggc	ranatatata	actaacegge	acactacaac	acctatcact	4560
ggtaaccacg	teteeeeac	geactatgtg	ataataaaa	acgctgcagc	ataataaa	4620
cagatectet	ctagtettae	cattacttag	tractora	ggcttcacca	ttaatataa	4680
gaggactgct	ccacgccatg	eceeggeceg	Lygoraagag	atgtttggga	attacacaca	4740
acggtgttga	ctgatttcaa	gacctggctc	cagtecaage	teetgeegeg	accyccygya	4800
gtccccttct	tctcatgtca	acgrgggtac	aagggagtet	ggcggggcga	ttaattaaa	4860
caaaccacct	gcccatgtgg	agcacagatc	accggacatg	tgaaaaacgg	trecargagg	4920
atcgtggggc	ctaggacctg	tagtaacacg	tggcatggaa	cattccccat	taacgcgtac	
accacgggcc	cctgcacgcc	ctccccggcg	ccaaattatt	ctagggcgct	graacaaara	4980
gctgctgagg	agtacgtgga	ggttacgcgg	gtgggggatt	tccactacgt	gacgggcatg	5040
accactgaca	acgtaaagtg	cccgtgtcag	gttccggccc	ccgaattctt	cacagaagtg	5100
gatggggtgc	ggttgcacag	gtacgctcca	gcgtgcaaac	ccctcctacg	ggaggaggtc	5160
acattcctqq	tcgggctcaa	tcaatacctg	gttgggtcac	agctcccatg	cgagcccgaa	5220
ccggacgtag	cagtgctcac	ttccatgctc	accgacccct	cccacattac	ggcggagacg	5280
gctaagcgta	ggctggccag	gggatctccc	ccctccttgg	ccagctcatc	agctatccag	5340
ctatctacac	cttccttgaa	ggcaacatgc	actacccgtc	atgactcccc	ggacgctgac	5400
ctcatcgagg	ccaacctcct	gtggcggcag	gagatgggcg	ggaacatcac	ccgcgtggag	5460
tcagaaaata	aggtagtaat	tttggagtct	ttcgagccgc	tccaagcgga	ggaggatgag	5520
agggaagtat	ccattccaac	ggagatcctq	cggaggtcca	ggaaattccc	tcgagcgatg	5580
cccatatooo	Gaccccaa	ttacaaccct	ccactgttag	agtcctggaa	ggacccggac	5640

taggtggtg	cagtggtaca	caaatataca	ttaccaccta	ccaaggcccc	tccgatacca	5700
cacgeeeee	0090990000	-333-3000	t	contatette	taccttaaca	5760
cctccacgga	ggaagaggac	ggrrycaatg	tcagaatcta	cogogococo	cgccccggcg	
gaggt.cgcca	caaaqacctt	cadcadctcc	gaatcgtcgg	ccgtcgacag	cggcacggca	5820
gagezegeea		ataaaaaaa	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	gatggaggt	tgagtcgtac	5880
acggcctctc	ctgaccagcc	ccccgacgac	ggcgacgcgg	gaccegacge	0949009040	
tectecatec	cccccttga	qqqqqaqccq	ggggatcccg	atctcagcga	cgggtcttgg	5940
200200005		tagtgaggag	stastatact	actcastata	ctacacatgg	6000
tetacegtaa	gcgaggaggc	Lagigaggac	gregeeger	goodgoogoo		
adadddddd	tgatcacgcc	atqcqctqcq	gaggaaacca	agctgcccat	caargcacrg	6060
~~~	bbasetan	aasassatta	atatataata	caacatctco	cadedeaade	6120
agcaactctt	tgctccgtca	Ccacaactty	giccalgeca	·	cagogoaago	
rtaraaraaa	agaaggtcac	ctttqacaqa	ctgcaggtcc	rggacgacca	ctacegggac	6180
	agatgaaggc	assagaat aa	acacttaacc	ctaaacttct	atccatagaa	6240
gtgctcaagg	agatgaagge	gaaggegeee	acagecaagg			6300
daadcctdta	agctgacgcc	cccacattcq	gccagatcta	aatttggcta	Lggggcaaag	
	acctatccag	assaggggt t	aaccacatco	actecatata	gaaggacttg	6360
gaegreegga	acctatccag	caaggeegee	aaccacacco	3000030303		6420
ctqqaaqaca	ctgagacacc	aattgacacc	accatcatgg	caaaaaacya	ggttttetge	
	agaagggggg	ccccaaccca	actcacctta	tcgtattccc	agatttgggg	6480
greeaaceag	agaaggggg	cegeaageea	900090000		taraaaata	6540
gttcgtgtgt	gcgagaaaat	ggccctttac	gatgtggtet	Ceacectece	ccaggccgcg	
staggtatt	catacggatt	ccaatactct	cctggacagc	gggtcgagtt	cctqqtqaat	6600
argggerere	cacacggacc			222 2 2	ttttaactca	6660
gcctggaaag	cgaagaaatg	ccctatgggc	ttegeatatg	acacecycig	LLLLGactca	
aggataacta	agaatgacat	ccatattaaa	gagtcaatct	accaatgttg	tgacttggcc	6720
acggccaccg	uguauguau		2-2-4-4444	tttagatggg	agarararta	6780
cccgaagcca	gacaggccat	aaggtegete	acagagegge	LLLacaregg	gggccccccg	
actaattcta	aagggcagaa	ctacaactat	caccaatacc	gcgcgagcgg	tgtactgacg	6840
200220000		anantattaa	+tanagacaa	ctacaaceta	traaactaca	6900
accagctgcg	gtaataccct	Cacatyttac	Ligaaggiceg	cracaaccra	cogugoogog	
. aageteeagg	actgcacgat	gctcgtatgc	ggagacgacc	ttgtcgttat	ctgtgaaagc	6960
aagoooagg	aagaggacga	2-2-2-2-2	agagaatta	cagaaactat	gactagatac	7020
gcggggaccc	aagaggacga	ggegageeta	egggeeeea	caaraaccar	gaocagaoao	•
tatacacaca	ctggggaccc	qcccaaacca	gaatacgact	tggagttgat	aacatcatgc	7080
b b a - b	tgtcagtcgc	adadda+dda	teteecaaaa	agatatacta	tctcacccqt	7140
teeteeaatg	Lgreagrege	gcacgargca	CCCggcaaaa	999090000		7200
gaccccacca	cccccttgc	gegggetgeg	tgggagacag	ctagacacac	tecagteaat	
testeration	gcaacatcat	catotatoco	cccaccttqt	gggcaaggat	gatectgatg	7260
ceerggerag	gcaacaccac	catgtatgeg		33300033500	5-2++-+	7320
actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaaageeet	agactyctag	
atatagggg	cctgttactc	cattgaggga	cttgacctac	ctcagatcat	tcaacgactc	7380
acctacgggg	cccgccaccc	caccagagooa		-t	taggataggt	7440
catggcctta	gcgcattttc	actccatagt	Lactercag	grgagarcaa	cagggcggcc	* **
tratorctra	ggaaacttgg	aataccaccc	ttqcqaqtct	ggagacatcg	ggccagaagt	7500
ccaegeceea	354444-55	222222	aggattaga	cttataacaa	gtacctcttc	7560
greegegera	ggctactgtc	ccaygggggg	agggccgcca	cccgcggcaa	900000000	
aactgggcag	taaggaccaa	qctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	7620
~~~5		-				
	~~++aa++aa	taattaaaaa	addadadaca	tatatcacag	cctatctcat	7680
ttatccagct	aattcattac	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	
acccaacccc	ggttcgttgc gctggttcat	tggttacagc gtggtgccta	gggggagaca ctcctacttt	tatatcacag ctgtaggggt	aggcatctat	7740
acccaacccc	ggttcgttgc gctggttcat	tggttacagc gtggtgccta	gggggagaca ctcctacttt	tatatcacag ctgtaggggt	aggcatctat	
gcccgacccc ctactcccca	ggttcgttgc gctggttcat accgatgaac	tggttacagc gtggtgccta ggggacctaa	gggggagaca ctcctacttt acactccagg	ctgtaggggt ccaataggcc	aggcatctat atcctgtttt	7740 7800
gcccgacccc ctactcccca tttccctttt	ggttcgttgc gctggttcat accgatgaac ttttttctt	tggttacagc gtggtgccta ggggacctaa ttttttttt	gggggagaca ctcctacttt acactccagg ttttttttt	tatatcacag ctgtaggggt ccaataggcc ttttttttt	aggcatctat atcctgtttt ttctcctttt	7740 7800 7860
gcccgacccc ctactcccca tttccctttt	ggttcgttgc gctggttcat accgatgaac ttttttctt	tggttacagc gtggtgccta ggggacctaa ttttttttt	gggggagaca ctcctacttt acactccagg ttttttttt	tatatcacag ctgtaggggt ccaataggcc ttttttttt	aggcatctat atcctgtttt ttctcctttt	7740 7800
gcccgacccc ctactcccca tttccctttt ttttcctct	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt	tggttacagc gtggtgccta ggggacctaa tttttttt ttctttcctt	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc	aggcatctat atcctgtttt ttctcctttt tagtcacggc	7740 7800 7860 7920
gcccgacccc ctactcccca tttccctttt tttttcctct	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc	7740 7800 7860 7920 7980
gcccgacccc ctactcccca tttccctttt tttttcctct	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc	7740 7800 7860 7920
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt qtqqgaatac	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta	7740 7800 7860 7920 7980 8040
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatqtttga	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt	7740 7800 7860 7920 7980 8040 8100
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatqtttga	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt	7740 7800 7860 7920 7980 8040
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcaccgt	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg	7740 7800 7860 7920 7980 8040 8100 8160
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcacctq	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcaccgt gatgctgtag	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg	7740 7800 7860 7920 7980 8040 8100 8160 8220
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcaccgt gatgctgtag cattccgaca	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcaccgt gatgctgtag cattccgaca	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc	7740 7800 7860 7920 7980 8040 8100 8160 8220
gcccgacccc ctactcccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcaccgt gatgctgtag cattccgaca tttctatgcg	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag caccgttct	aggcatctat atcctgttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340
gcccgacccc ctactcccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa ccagtcctq	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcaccgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag cacccgttct tacttggagc	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8400
gcccgacccc ctactcccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa ccagtcctq	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcaccgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag cacccgttct tacttggagc	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340
gcccgacccc ctactcccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggccgccg	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa cccagtcctg acccqtcctq	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcaccgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc tggatcctct	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8400 8460
gcccgacccc ctactcccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct tacgcgatca	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggccgccg tggcgaccac	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcaccgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc tggatcctct qgcgcctata	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg tcgccgacat	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8460 8520
gcccgacccc ctactcccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct tacgcgatca	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggccgccg tggcgaccac	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcaccgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc tggatcctct qgcgcctata	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg tcgccgacat	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8400 8460
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct tacgcgatca gcatcaccg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg tgcggttgct cgqgctcatq	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcaccgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc tggatcctct ggcgcctata aqcgcttgtt	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg tcgccgacat tcgcggtggg	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatgggc tatggcc	7740 7800 7860 7920 7980 8040 8100 8220 8280 8340 8460 8520 8580
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct tacgcgatca ggaatcaccg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggccgccg tggcgaccac gcgccacagg ctcgccactt	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg tgcggttgct cgggctcatg qttgggcgc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagctcgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc tggatcctct ggcgcctata agcgcttgt atctccttgc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg tcgccgacat tcgcgacat tcgcgtggg atgcaccatt	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggtgca ccttgcggcg	7740 7800 7860 7920 7980 8040 8100 8220 8280 8340 8460 8520 8580 8640
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct tacgcgatca ggaatcaccg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggccgccg tggcgaccac gcgccacagg ctcgccactt	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg tgcggttgct cgggctcatg qttgggcgc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagctcgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc tggatcctct ggcgcctata agcgcttgt atctccttgc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gtatgaaatc gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg tcgccgacat tcgcgacat tcgcgtggg atgcaccatt	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggtgca ccttgcggcg	7740 7800 7860 7920 7980 8040 8100 8220 8280 8340 8460 8520 8580 8640 8700
gcccgacccc ctactcccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct tacgcgatca ggaatcaccg gaagatcggg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggccgccg tggcgaccac gcgccacagg ctcgccactt ccgggggact acggcctcaa	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg tgcggttgct cgggctcatg gttgggcgcc cctactactq	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagctcgt gatgctgtag cattccgaca tttctatgcg ctcgcttcgc tggatcctct ggcgcctata agcgcttgtt atctccttgc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggctt gcatcgccag cacccgttct tacttggagc acgccgacat tcgccgacat tcgcgtggg atgcaccatt taatgcagga	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc catcgtggcc catcgtggcc catcgtggcc catcgtggcc catcgtggca tcatggcggt	7740 7800 7860 7920 7980 8040 8100 8220 8280 8340 8460 8520 8580 8640 8700
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct tacgcgatca ggaatcaccg gaagatcggg gcgcccgtgg gcggtgctca	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggccgccg tggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatga	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgcttgac cgcgccacta tcatgtttga ctaacgcagt cgtcaccctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg tgcggttgct cgggctcatg cttgggcgcc cctactactg cttgagagcc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagctcgt gatgctgtag cattccgaca tttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg tcgccgacat tcgcgtggg atgcaccatt taatgcagga tcagctctt	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcg tatggcgg tatggcgg tatggtggca ccttgcggcg gtcgcataag ccggtgggcg	7740 7800 7860 7920 7980 8040 8100 8220 8280 8340 8400 8460 8520 8580 8640 8700
gccgaccc ctactccca tttcccttt ttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg ttcgaccgct tacgcgatca ggcatcaccg gaagatcggg gccccgtgg gcgctgctag	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcg	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg tgcggttgct cgggctcatg ttgcggcccctactactg cttgagagcc cctactactg cttgagagcc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagctcgt gatgctgtag cattccgaca tttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag actgtcttc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg tcgccgacat tcgcgtggg atgcaccatt taatgcagga tcagctctt taatgcagga tcagctctt	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcg gtatgcgg tatggcg actgcgataag ccttgcggcg gtcgcataag actcgtagga	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8460 8520 8580 8640 8700 8760 8820
gccgaccc ctactccca tttcccttt ttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg ttcgaccgct tacgcgatca ggcatcaccg gaagatcggg gccccgtgg gcgctgctag	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcg	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcaa cccagtcctg acccgtcctg tgcggttgct cgggctcatg ttgcggcccctactactg cttgagagcc cctactactg cttgagagcc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagctcgt gatgctgtag cattccgaca tttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag actgtcttc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggctt gcatcgccag cacccgttct tacttggagc acgccggacg tcgccgacat tcgcgtggg atgcaccatt taatgcagga tcagctctt taatgcagga tcagctctt	aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcg gtatgcgg tatggcg actgcgataag ccttgcggcg gtcgcataag actcgtagga	7740 7800 7860 7920 7980 8040 8100 8220 8280 8340 8400 8460 8520 8580 8640 8700
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg gtgctgctag tccgaccgct tacgcgatca ggaatcaccg gaagatcggg gcgctgctca ggagagcgtca ggagagcgtca cggggcatga	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcgc caggggact	tggttacagc gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcaa cccagtcctg tgcggttgct cgggctcatg tgcggctcatg gttgggcgc cctactactg cttgagagcc cctactactg cgcacttatg	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagctcgt gatgctgtag cattccgaca tttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag actgtcttcc ggcgaggacc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggctt gcatcgccag cacccgttct tacttggagc acgccgacat tcgccgacat tcgcgtggg atgcaccatt taatgcagga tcagctctt taatgcagga tcagctctt taatgcagga tcagctctt	aggcatctat aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggtggca ccttgcggcg gtcgcataag ccggtgggcg actcgtagga gagcgcgacg	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8460 8520 8580 8640 8700 8760 8820 8880
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggaatcaccg gaagatcggg gccccgtgg gcggtgctca ggagagcgtc cggggcatga caggtgccgg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcgc cagcgctttg	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcaa cccagtcctg tgcggttgct cgggctcatg tgcggctcatg gttgggcgcc cctactactg cttgagagcc cgtacttatg ggtcattttc ggtattcgg	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcttatca cagcaccgt gatgctgtag cattccgaca ttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag actgtcttcc ggcgaggacc atcttgc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggctt gcatcgccag cacccgttct tacttggagc acgccgacat tcgcggacat tcgcgacat tcgcgtggg atgcaccatt taatgcagga ccatttatcatgca gcttcatcatgca	aggcatctat aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcg gtcgcataag ccttgcggcg actcgtagga agccttcgtc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8460 8520 8580 8640 8700 8760 8820 8880
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggaatcaccg gaagatcggg gccccgtgg gcggtgctca ggagagcgtc cggggcatga caggtgccgg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcgc cagcgctttg	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcaa cccagtcctg tgcggttgct cgggctcatg tgcggctcatg gttgggcgcc cctactactg cttgagagcc cgtacttatg ggtcattttc ggtattcgg	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcttatca cagcaccgt gatgctgtag cattccgaca ttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag actgtcttcc ggcgaggacc atcttgc	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggctt gcatcgccag cacccgttct tacttggagc acgccgacat tcgcggacat tcgcgacat tcgcgtggg atgcaccatt taatgcagga ccatttatcatgca gcttcatcatgca	aggcatctat aggcatctat atcctgtttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcg gtcgcataag ccttgcggcg actcgtagga agccttcgtc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8460 8520 8580 8640 8700 8760 8820 8880
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggaatcaccg gaagatcggg gccccgtgg gcgctgcta ggagagcgtc cggggcatca cggggcatca caggtgcca	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcgc cagcgctctg tgtcgctctg	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcaa cccagtcctg tgcggttgct cgggctcatg tgcggctcatg gttgggcgcc cctactactg cttgagagcc cttgagagcc cttgagagcc cttgagagcc cttgagagcc cttactg tttgggttttc ggtattttc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcttatca cagcaccgt gatgctgtag cattccgaca ttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag actgtcttcc ggcgaggacc atcttgcacg acgcaggacca	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggctt gcatcgccag cacccgttct tacttggagc acgccgacat tcgccgacat tcgcgtggg atgcaccatt ttatcatgcag gcttccttatcatgcag tcagctcctt ttatcatgca gcttcgctg	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcg gtcgcataag ccttgcggcg actcgtaggac agcctcgtcgcc catggcgcc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcaggg gccccgtgg gcggtgctca ggagagcgtc cggggcatga caggtgccgg atgatcggg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcgc cagcgctctg tgtcgcttgc cagcgctctg	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcaa cccagtcctg tgcggttgct cgggctcatg tgcggctcatg gttgggcgcc cctactactg cttgagagcc cgtacttatg ggtcattttc ggtattcgga gttgcgcacttatg ggtcattttc ggtattcgga tttcggcgag gctggcgtc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcttatca cagcaccgt gatgctgtag cattccgaca ttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag actgtcttcc ggcgaggacc atcttgcacg acgcaggag	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccgacat tcgcggacat tcgcgcgacat tcagccatt ttatcatgca gcttccct tcattgcagg atgcaccatt ttatcatgca gcttcgctg ccctcgctca ttatcgcgg gctggatgg	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcg gtcgcataag ccttgcggcg actcgtaggac agccttcgtc catggcgcc catggcgcc cctcgtcgtc ccttccccatt	7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcaggg gccccgtgg gcggtgctca ggagagcgtc cggggcatga caggtgccgg atgatcggg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcctcttgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcgc cagcgctctg tgtcgcttgc cagcgctctg	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcaa cccagtcctg tgcggttgct cgggctcatg tgcggctcatg gttgggcgcc cctactactg cttgagagcc cgtacttatg ggtcattttc ggtattcgga gttgcgcacttatg ggtcattttc ggtattcgga tttcggcgag gctggcgtc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcttatca cagcaccgt gatgctgtag cattccgaca ttctatgcg ctggatcctct ggcgcctata agcgcttgt atctccttgc ggctgcttcc ttcaacccag actgtcttcc ggcgaggacc atcttgcacg acgcaggag	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccgacat tcgcggacat tcgcgcgacat tcagccatt ttatcatgca gcttccct tcattgcagg atgcaccatt ttatcatgca gcttcgctg ccctcgctca ttatcgcgg gctggatgg	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcg gtcgcataag ccttgcggcg actcgtaggac agccttcgtc catggcgcc catggcgcc cctcgtcgtc ccttccccatt	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcagg gccccgtgg gcggtgctca ggagagcgtc cggggcatga caggtgcccg atgatcggc atgatcggc	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgc ctatcgtcgc cagcgctctg tgtcgcttgc cagcgctttgc cagcgctttgc ctatcgtcgc tgtcgctttgc	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgctacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcat accegtcctg tgcggttgct cgggctcat gttgggcgc cctactactg gttgagagc cttgagagcc cttgagagcc cttgagagcc cttgagagcc cgcacttatg ggtcattttc ggtattcgga tttcggcgag gctggcgttc cggcatcggg	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcttatca cagcttatcg cattccgaca tttctatgcg ctcgatcctct ggcgctata agcgcttat atctccttgc ggctgcttcc ttcaacccag actgtctcc ggcgaggacc atcttgcacg actgtctcc ggcgaggacc atcttgcacg actgcacgagacca gcgacgcgag atgcccgcg	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccgacat tcgcggacat tcgcgacat tcagctcat taatgcagga tcaccatt ttatcatgca gcttcgcta tcagctcat ttatcatgca gcttcgcta tcagccat tcagctcat tcagcacat tcagcacat tcagcacat tcagcacat tcagcacat tcagcacat tcagcacac tcag	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cggagcactg cactatcgac catcgtggcc caccgatggg tatggcgg tatggcgg tatggcgg tatggcgg tatggtggca ccttgcgcataag ccgtaggcg actcgtagga agccttcgtc catggcggcc catggcgacg agccttcgtc catggcggcc cttccccatt gctgtcagg	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcagg gcgctgctag ccggtgctca ggagagcgtc cggggcatga caggtgccgg atgatcggc actggtcccg atgatcgcg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gctattgcg cgctatatgc ttggcgaccac gcgccacagg ctcggggact acggcctcaa gaccgatgcc ctatcgtcgc cagcgctttgc cagcgctttgc cagcgctttgc cagcgctctg tgtcgcttgc acggctttgc acggctttgc cagcgctctg acggctttgc ccaccaaacg gctacgtctt tcgcttccgg acgaccatca	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcat acccgtcctg tgcggttgct cgggctcat gttgggccc cttactactg gttgggccc cttgagagcc cttgagagcc cttgagagcc cttgagagcc cgcacttatg ggtcattttc ggtattcgga gctggcgtt cggcatcggg gctggcgtt cggcatcggg	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcttatca cagcttatcg cattccgaca ttctatgcg ctggatcctct ggcgctata agcgcttat atctccttgc ggcttccttgc tgatcctct ggcgctata actgcttcc ttcaacccag actgtcttcc ttcaacccag actgtctcc ggcgaggacc atcttgcacg actgcacgcgac atgcacgcgac caaggacca	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccgacat tcgcggacat tcgcgcgacat tcagctcat taatgcagga tcagctct taatgcagga tcagctcat ttatcatgca gcttcgctc ttatcatgca gcttcgctc ttatcagct tcagctcat	aggcatctat atcctgtttt ttctcctttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cactatcgac catcgtggcc caccgatggg tatggtggca ccttgcgcg gtcgtaggcg accgtagga gcgtaggcg accgtagga ccggtggcc catggcgacg accgtagga gcgtaggacg accgtagga gagccttcgtc catggcgcc catggcgcc catggcgacg agccttcgtc catggcgcc cttccccatt gctgtccagg taccagccta	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcagg gcgctgctag ccggtgctca ggagagcgtc cggggcatga caggtgccgg atgatcggc actggtcccg atgatcgcg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gctattgcg cgctatatgc ttggcgaccac gcgccacagg ctcggggact acggcctcaa gaccgatgcc ctatcgtcgc cagcgctttgc cagcgctttgc cagcgctttgc cagcgctctg tgtcgcttgc acggctttgc acggctttgc cagcgctctg acggctttgc ccaccaaacg gctacgtctt tcgcttccgg acgaccatca	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcat acccgtcctg tgcggttgct cgggctcat gttgggccc cttactactg gttgggccc cttgagagcc cttgagagcc cttgagagcc cttgagagcc cgcacttatg ggtcattttc ggtattcgga gctggcgtt cggcatcggg gctggcgtt cggcatcggg	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca cagcttatca cagcttatcg cattccgaca ttctatgcg ctggatcctct ggcgctata agcgcttat atctccttgc ggcttccttgc tgatcctct ggcgctata actgcttcc ttcaacccag actgtcttcc ttcaacccag actgtctcc ggcgaggacc atcttgcacg actgcacgcgac atgcacgcgac caaggacca	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccgacat tcgcggacat tcgcgcgacat tcagctcat taatgcagga tcagctct taatgcagga tcagctcat ttatcatgca gcttcgctc ttatcatgca gcttcgctc ttatcagct tcagctcat	aggcatctat atcctgtttt ttctcctttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cactatcgac catcgtggcc caccgatggg tatggtggca ccttgcgcg gtcgtaggcg accgtagga gcgtaggcg accgtagga ccggtggcc catggcgacg accgtagga gcgtaggacg accgtagga gagccttcgtc catggcgcc catggcgcc catggcgacg agccttcgtc catggcgcc cttccccatt gctgtccagg taccagccta	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180 9240
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcagg gcgctgctag ccgggtgctca ggagagcgtc caggtgcccg atgatcggc actggtcccg atgatcgccg atgatcgccg atgatcgccg atgatcaccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actgctcccg actgctccccg actgctccccg actgctccccg actgctccccg actgctcccg actgctccccg actgctcccccccccc	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gctattgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgact acggcgtcaa gaccgatgcc ctatcgtcgc cagcgctctg tgtcgcttgc ccaccaaacg ctaccttgcg ccaccaaacg ctgcccacac tcgcgctctcac ccaccaaacg	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcat acccgtcctg tgcggttgct cgggctcat tcatgttga cccagtcctg ggatatcgtc gttgatgcc cctactact gttgggccc cttgagagcc cttgagagcc cgcacttatc ggtattcggg tttcggcgag gctggcgttc ggtatcggg	ggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcaccgt gatgctgtag cattcctatgcg ctctctatgcg ctcgatcctct ggcgcttat acccaga atcctcttgc ggcttccttgc tgcaccagagacca actgtcttcc ttcaacccag actgtctcc actgtctcc ttcaacccag actgcaggacca gcgacgcgag atgcccgcgt caaggatcgc gcgatttatg gcgattat	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccgacat tcgcggacat tcgcgcgacat tcagctcat tcatcatgca gcttcatcatgcac atgcaccatt tcatcatgca gcttcactcaccc tcactcgctca tcacccgttct tcactcgctca tcacccgttct tcactcgctca tcacccgtcac tcacccgctcac tcacccggacac tcacccgctcac tcacccggacac tcacccgctcac tcacccggaccac tcacccggaccac tcacccgcccc	aggcatctat atcctgtttt ttctcctttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg gttatgccg tcactatggc cactatcgac catcgtggcc caccgatggg tatggtggca ccttgcgcg gtcgtaggcg accgtaggcg accgtaggcg accgtaggcg accgtaggcg accgtagga gagcgcact catggcgcc catggcgacg agccttcgtc catggcggcc catggcgacg agccttcgtc catggcggcc cttccccatt gctgtccagg taccagccta gagcacatgg	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcagg gcgctgctag ccgggtgctca ggagagcgtc cggggcatga caggtgccgg atgatcggc actggtcccg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gctattgcg cgctatatgc ttggcgaccac gcgccacagg ctcggcgacta acggcctcaa gaccgatgcc ctatcgtcgc cagcgctttgc cagcgctttgc cagcgctttgc ccaccaaacg ctcgcttgc ccaccaaacg	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcat acccggtccatg tgcggctcatg tgcggctcatg tgcggctcatg gttgggccc cttgagagcc cttactactg ggttgct ggtcattttc ggtattcgg ggtcattttc ggtattcgga tttcggcgag gctgccgcc gggacagctt gatcgccgcc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcatcatc gatgctgtag cattcctatgc ctcgatctct ggcgcttat acctcttgc ggcgcttat acctcttgc ggcgcttcc ttcaacccag actgtcttc ggcgaggacc actgtcttct ggcgaggacc actgtctct ggcgaggacc actgtctct ggcgaggacc actgtctct ggcgaggacc actgtctct ctaacccag actgtctct ggcgaggacc actgtctct ctaacccag actgtctct ctaacccag ctataccttg ctataccttg	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccgacat tcgcggacat tcgcgcgacat tcagctcat tcatcatgca gcttccct tactgcgg atgcaccatt tcatcatgca gcttcgctc tcactcgctc tcactcgctc tcactcgctc tcactcgctc tcactcgctca tcacccgggacat tcaccccg tcacccggacat tcaccccc tcacccggacac tcaccccgctca tcaccccggacac tcaccccgctca tcacccccc tcacccccccccc	aggcatctat atcctgtttt ttctcctttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg gttatgccg tcactatggc cactatcgac catcgtggcc catcgtggcc catcgtggcg tatggtggca ccttgcgcg gtcgtaggcg acctgtagga gcgtaggcg acctcgtc catggcgcc catggcgcc catggcgcc catggcgcc agccttcgtc catggcggcc catggcggcc catggcggcc catggcgcc catggcggcc catggcggcc catggcggcc catggcggcc catggcggcc catggcgcc catggcggcc catggcgcc catggcggcc catggcggcc catggcggcc catggcgcc catggcggcc catggcgcc catggc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180 9300
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcagg gcgctgctag ccgggtgctca ggagagcgtc cggggcatga caggtgccgg atgatcggc actggtcccg actgcccg acggtcccg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gctatatgc ttggcgaccac ttggcgaccac gcgccacagt ccgggggact acggcctcaa gaccgatgc ctatcgtcgc cagcgctttgc cagcgctttgc ccaccaaacg gctacgtttccgg ccaccaacg ctatcgtcgc cagcgctctg tgtcgcttgc ccaccaaacg gctacgttt	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcat acccggtcat tcatgtttga cctaacgcagt cgtcacctg ggatatcgtc gttgatgcc cctactatg ggttgggcc cctactatg ggttgggcc cctactatg ggtcatttt ggtattcgga tttcggcgag gctggcgct cggcatcggc cacctcgacc cacctcgacc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcatcatc gatgctgtag cattcctatgcg ctctatatgcg ctcgatcctct ggcgcttat acctctatgc gctgctctc ggcgcttat acctcttgc ggcgcttat acctcttgc ggcgcttcc ttcaacccag actgcagagacc actgtctct ggcgaggacc actgtctct ggcgaggacc actgtctct tgaacgcgag caaggatcgc gcgatttatg ctataccttg tgaatgaag	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccggacat tcgcggacat tcgcggacat tcagctcat tcatcatgca gcttcatcatgcaccatt tcatcatgca gcttcaccct tcatcaccc tcatcaccc tcatcaccc tcatcaccc tcatcaccc ccgcctcac tcacccggacac tcacccggacac tcacccgcc ccgcctcac tcacccggacac tcgcggacac tcgcggcac	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cactatcgac catcgtggcc catcgtggcc catcgtggca cattggcgg tatggcgg tatggcgg tatggcg tatggtggca ccttgcgcataag ccggtgggcg actcgtagga gacctcgtc catggcgcc catggcgacg agccttcgtc catggcggcc cttccccatt gctgtccagg taccagccta gagcacatgg ccgcttacgc ctcgctaacg ccgcttacgc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180 9360
gccgacccc ctactccca tttcccttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggatcaccg gaagatcagg gcgctgctag ccgggtgctca ggagagcgtc cggggcatga caggtgccgg atgatcggc actggtcccg actgcccg acggtcccg	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gctatatgc ttggcgaccac ttggcgaccac gcgccacagt ccgggggact acggcctcaa gaccgatgc ctatcgtcgc cagcgctttgc cagcgctttgc ccaccaaacg gctacgtttccgg ccaccaacg ctatcgtcgc cagcgctctg tgtcgcttgc ccaccaaacg gctacgttt	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcat acccggtcat tcatgtttga cctaacgcagt cgtcacctg ggatatcgtc gttgatgcc cctactatg ggttgggcc cctactatg ggttgggcc cctactatg ggtcatttt ggtattcgga tttcggcgag gctggcgct cggcatcggc cacctcgacc cacctcgacc	gggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcatcatc gatgctgtag cattcctatgcg ctctatatgcg ctcgatcctct ggcgcttat acctctatgc gctgctctc ggcgcttat acctcttgc ggcgcttat acctcttgc ggcgcttcc ttcaacccag actgcagagacc actgtctct ggcgaggacc actgtctct ggcgaggacc actgtctct tgaacgcgag caaggatcgc gcgatttatg ctataccttg tgaatgaag	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc acgccggacat tcgcggacat tcgcggacat tcagctcat tcatcatgca gcttcatcatgcaccatt tcatcatgca gcttcaccct tcatcaccc tcatcaccc tcatcaccc tcatcaccc tcatcaccc ccgcctcac tcacccggacac tcacccggacac tcacccgcc ccgcctcac tcacccggacac tcgcggacac tcgcggcac	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg ggttatgccg tcactatggc cactatcgac catcgtggcc catcgtggcc catcgtggca cattggcgg tatggcgg tatggcgg tatggcg tatggtggca ccttgcgcataag ccggtgggcg actcgtagga gacctcgtc catggcgcc catggcgacg agccttcgtc catggcggcc cttccccatt gctgtccagg taccagccta gagcacatgg ccgcttacgc ctcgctaacg ccgcttacgc	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180 9300
gccgacccc ctactccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggagatcacg gaagatcagg gcgctgctag ccgggtgctca ggagagcgtc caggtgccgg atgatcggc actggtccg actggtcccg actgcgctcac actggtcccg actgcgctccac actggtcccg actgctcccg actgcccg actgcccg actgcccg actgccccg actgccccc actgcccccg actgccccc actgccccc actgccccc actgccccc actgccccc actgccccc actgccccc actgccccc actgcccccc actgcccccc actgcccccc actgcccccc actgcccccc actgcccccc actgcccccc actgcccccc actgcccccc actgccccccc actgccccccc actgccccccc actgccccccc actgcccccccc actgccccccccc actgccccccccc actgccccccccc actgcccccccccc	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gctatatgc ttggcgaccac ttggcgaccac tcgggggact acggcctcaa gaccgatgc ctatcgtcgc cagcgctctg tgtcgcttgc cagcgctttgc cagcgctttgc ccaccaaacg gctacgtttccgg tcgctacgt tcgcttcgc tcgcttcgc tcaccaaacg gctacgtctt tcgcttccgg acgaccatca ctggaccgct tcgctcaa	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcat acccggtcat tcatgttga ccagtcat ggttgatgcc gttgatgcc cctactag ggttgggcc cctactatg ggttgggcc cctactatg ggtcatttt ggtattcgga gctggcgat ggtcatcgga gctggcgat cggcatcgga gatcgccac ggagccaatc	ggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcaccgt gatcctct gcatctct ggcgcttata accggcttata accggcttata accggcttata accgcttata gcgcttata accgcttata gcgcttata accgccttata accgccttata accgccttata accgccttata accgcagagacca actgcaccaga actgcaccaga accgcagagacca gcgaccgcgagacca gcgatttatg ctataccttg ctataccttg gcgatttatg accgcagagacca accgcagacca accgcagacca accaggatcca gcgatttatg ctataccttg ctataccttg	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc tcgccgacat tcgcggacat tcgcgcgacat tcagctcat tcagctcat tcagctcat tcagctcat ccgcgtcat tcagctcac tcagctcac tcagctcac tcagcacat tcagcacat tcagcacat tcagcacat tcagcacat tcagcacac cccgacaca gctggatggc tcagcacac tcagcacac gcagacac tcagcacac tcagcacac gcagacac tcagcacac tcagcacac gcagacac tcagcacac tcagcacac tcagcacac tcagcacac tcagcacac tcagcacac gcagacac gagaactgt	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt ttaatgcgg ttaatgccg tcactatggc cggatgcactg cactatcgac catcgtggcc caccgatgggca cattggcgg tatggcgg tatggcgg tatggcgc catcgtggcc catcgtggcc ccttgcgcatagg agccttcgtc catggcggcc catggcgacg agccttcgtc catggcggcc cttccccatt gctgtccagg taccagccta gagcacatgg cgcgttgcgt catggcgaca	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180 9360 9420
gccgacccc ctactccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggagatcacg gaagatcagg gcgctgctag ccgggtgctca ggagagcgtc caggtgcccg atgatcggc actggtcccg actggtcccg actggtcccg gacgcgctca caggtgcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actgcgctca caggttccccg actgcgctcca caggttccccg actcccccccccc	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccagg gacccaattc agttaaattg tcctgggcac gctattgcg tggcgaccac ttggcgaccac ttggcgaccac ttggcgaccac tcgggggact acggcctcaa gaccgatgc catcgtcgc catcgtcgc tgtcgcttgc catcgtcgc tgtcgcttgc ccaccaaacg gctacgtttccgg tgtcgcttgc ccaccaaacg tcgcttcgc ccaccaaacg gctacgtttccgg tgtcgcttgc ccaccaaacg gctacgttt	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcct accagtcctg tgcggctcat tcatgttga cccagtcctg ggatatcgtc ggttgatgcc cctactactg ggttgggcc cctactactg ggttgggcc cctactactg ggtcatttt ggtattcgga gctggcgac gctggcgac cggacagct gatcgcgcc cacctcgacc ggagccaatc tccatcggt	ggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcatcatc gatgctgtag cattcctatgc ctcgatctct ggcgcttat acactcag ctctattgc ctcgatctct ggcgcttat accttgc tggatcctct ggcgcttat acctgcttcc ttcaacccag actgcagagacc actgtctct ggcgaggacc actgcaggacca gcgacgcgag caaggatcgc gcgatttatg ctataccttg tgaatgccc ccgccatct	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc tcgccgacat tcgcggacat tcgcgcacat tcagctgcacat tcagctcat tcagctcaccc tcatcagccat tcagcagcact ccgcctcag ccctcgctca tcatcagcac gctggatggc tcagcacat tcagcagcac tcagcagcac tcagcagcac tcagcagcac tcagcagcac tcagcagcac tcagcagcac ccggaacac ccgagaacc ccgagaacc ccgagaacc ccgagaacc ccgagaacc ccgagaacc ccgagaacc ccagcagcac ccagcagcac	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt ttaatgcgg ttaatgccg tcactatggc cactatcgac catcgtggcc catcgtggcc catcgtggca tatggtggca ccttgcgcataag gcttgcgtagga gcgtaggcg actcgtagga gcgtaggcg actcgtcgtc catggcgcc catggcgca agccttcgtc catggcggcc catggcgacg agccttcgtc catggcggcc cttccccatt gctgtcaag cgcgttgcgt taccagccta gagcacatgg cacagcgca aagcgcaaa acgcggcgca	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 9000 9120 9180 9360 9420 9480
gccgacccc ctactccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggagatcacg gaagatcagg gcgctgctag ccgggtgctca ggagagcgtc caggtgcccg atgatcggc actggtcccg actggtcccg actggtcccg gacgcgctca caggtgcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actgcgctca caggttccccg actgcgctcca caggttccccg actcccccccccc	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccagg gacccaattc agttaaattg tcctgggcac gctattgcg tggcgaccac ttggcgaccac ttggcgaccac ttggcgaccac tcgggggact acggcctcaa gaccgatgc catcgtcgc catcgtcgc tgtcgcttgc catcgtcgc tgtcgcttgc ccaccaaacg gctacgtttccgg tgtcgcttgc ccaccaaacg tcgcttcgc ccaccaaacg gctacgtttccgg tgtcgcttgc ccaccaaacg gctacgttt	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcct accagtcctg tgcggctcat tcatgttga cccagtcctg ggatatcgtc ggttgatgcc cctactactg ggttgggcc cctactactg ggttgggcc cctactactg ggtcatttt ggtattcgga gctggcgac gctggcgac cggacagct gatcgcgcc cacctcgacc ggagccaatc tccatcggt	ggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcatcatc gatgctgtag cattcctatgc ctcgatctct ggcgcttat acactcag ctctattgc ctcgatctct ggcgcttat accttgc tggatcctct ggcgcttat acctgcttcc ttcaacccag actgcagagacc actgtctct ggcgaggacc actgcaggacca gcgacgcgag caaggatcgc gcgatttatg ctataccttg tgaatgccc ccgccatct	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc tcgccgacat tcgcggacat tcgcgcacat tcagctgcacat tcagctcat tcagctcaccc tcatcagccat tcagcagcact ccgcctcag ccctcgctca tcatcagcac gctggatggc tcagcacat tcagcagcac tcagcagcac tcagcagcac tcagcagcac tcagcagcac tcagcagcac tcagcagcac ccggaacac ccgagaacc ccgagaacc ccgagaacc ccgagaacc ccgagaacc ccgagaacc ccgagaacc ccagcagcac ccagcagcac	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt ttaatgcgg ttaatgccg tcactatggc cactatcgac catcgtggcc catcgtggcc catcgtggca tatggtggca ccttgcgcataag gcttgcgtagga gcgtaggcg actcgtagga gcgtaggcg actcgtcgtc catggcgcc catggcgca agccttcgtc catggcggcc catggcgacg agccttcgtc catggcggcc cttccccatt gctgtcaag cgcgttgcgt taccagccta gagcacatgg cacagcgca aagcgcaaa acgcggcgca	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180 9360 9360 9420
gccgacccc ctactccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggagatcacg gaagatcagg gcgctgctag ccgagagacgtca ggagagcgtca ggagagcgtca caggtgccgg atgatcggc actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg caggtgcccg tcaggtcccg actggtcccg actggtcccg actgcgcccg actgcccg actgcccg actgcccg actgcccg actgcccg actgcccg actgcccg actgcccc ccaggtcccc ccaggtcccc ccaggtcccc ccacccttcc ccaacccttcc tccc	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcgctatatgc ttggcgaccac ttggcgaccac tcgggggact acggcctcaa gaccgatgc ctatcgtcgc cagcgctctg tgtcgcttgc cagcgctctg tgtcgcttgc ccaccaaacg gctacgtct tcgcttccgg tgtcgcttgc ccaccaaacg gctacgtct tcgcttccgg acgaccata ccggaccatc ccaccaaacg gctacgtct tcgcttccgg acgaccatca ccggaccgct tcgcttccgg acgaccatca ccggaccgct catggaccatca ccgaccgct catggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccacacac ccatggaccacac ccacacacac ccatggaccacac ccacacacac ccatggaccacac ccacacacac ccacacacac ccacacaca	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcct accagtcctg tgcggctcat tcatgttga cccagtcctg ggatatcgtc cgtggctcat gttgggccc cttgaggcc cttactggg ttgggcgc cttattgg ggttgct ggtcatttt ggtattcgg ggtcatttt ggtattcgga tcggcgcc cgcactcgga gctggccc cacctcgacc ggagccaatc tccatcggg	ggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcttatca caggcactgt gatgctgtag cattcctatgcg ctctattgcg ctcgatcctct ggcgcttat acctggcttcc tggatcctct ggcgcttat acctggcttcc tgatcctct ggcgcttat acctgcttcc ttcaacccag actgcaggacc actgcaggacca gcgacgcgag caaggatcac gcgatttatg ctataccttg tgaatcctt tgaatgcaca caggatcct tgaatgcaca caggatcct tgaatgcaca tcataccttg ctataccttg ctataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc tcgccgacat tcgcggacat tcgcgcacat tcagccgttc ttatcatgca gctgcacat tcagccgtcat tcagctcac ccgcctcag cctcgctca tcagcagccat tcagcagcac tcagcagcac tcagcagcac tcagcagcac tcagcagcac ccgcacagcac cagcagcac cagcagcac cagcagcac	aggcatctat atcctgtttt ttctcctttt ttctcctttt tagtcacggc gcctctctgc ccgcgtttta ttaatgcggt taacaatgcg gttatgccg tcactatggc cactatcgac catcgtggcc catcgtggca catcgtggca catcgtggca catcgtggca catcgtggca catcgtggca catcgtggca catcgtggca ccttgcgcataag agccttcgtc catggcgacg agccttcgtc catggcggca agccttcgtc catggcggca catggcgca agcgcgcata gcgcgttgcc ctccccatt gctgcaaa agcgcgcaa acgcggcgca acgcgttgcgt ctcgctaacg catggcgca catggcgca catggcgca catggcgca catgcgcaaa acgcggcgca acgcgttgcgt ctcgctaacg cgcttgcgt ctcgctaacg cgcgttgcgt ctcgctaacg	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180 9360 9420 9480 9540
gccgacccc ctactccca tttccctttt tttttcctct tagctgtgaa agatcaagta gcatattgac agtttatcac ctcatcgtca gtactgccgg tccgaccgct tacgcgatca ggagatcacg gaagatcagg gcgctgctag ccgagagacgtca ggagagcgtca ggagagcgtca caggtgccgg atgatcggc actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg actggtcccg caggtgcccg tcaggtcccg actggtcccg actggtcccg actgcgcccg actgcccg actgcccg actgcccg actgcccg actgcccg actgcccg actgcccg actgcccc ccaggtcccc ccaggtcccc ccaggtcccc ccacccttcc ccaacccttcc tccc	ggttcgttgc gctggttcat accgatgaac ttttttctt tttttcctt aggtccgtga ctcctgcagg gacccaattc agttaaattg tcctcggcac gcgctatatgc ttggcgaccac ttggcgaccac tcgggggact acggcctcaa gaccgatgc ctatcgtcgc cagcgctctg tgtcgcttgc cagcgctctg tgtcgcttgc ccaccaaacg gctacgtct tcgcttccgg tgtcgcttgc ccaccaaacg gctacgtct tcgcttccgg acgaccata ccggaccatc ccaccaaacg gctacgtct tcgcttccgg acgaccatca ccggaccgct tcgcttccgg acgaccatca ccggaccgct catggaccatca ccgaccgct catggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccgct ccatggaccatca ccgaccacacac ccatggaccacac ccacacacac ccatggaccacac ccacacacac ccatggaccacac ccacacacac ccacacacac ccacacaca	tggttacagc gtggtgccta gtggtgccta ggggacctaa ttttttttt ttctttcctt gccgccacta tcatgtttga ctaacgcagt cgtcacctg ggatatcgtc gttgatgcct accagtcctg tgcggctcat tcatgttga cccagtcctg ggatatcgtc cgtggctcat gttgggccc cttgaggcc cttactggg ttgggcgc cttattgg ggttgct ggtcatttt ggtattcgg ggtcatttt ggtattcgga tcggcgcc cgcactcgga gctggccc cacctcgacc ggagccaatc tccatcggg	ggggagaca ctcctacttt acactccagg ttttttttt tggtggctcc tgcagagagt gtgggaatac cagcttatca caggcttatca caggcactgt gatgctgtag cattcctatgcg ctctattgcg ctcgatcctct ggcgcttat acctggcttcc tggatcctct ggcgcttat acctggcttcc tgatcctct ggcgcttat acctgcttcc ttcaacccag actgcaggacc actgcaggacca gcgacgcgag caaggatcac gcgatttatg ctataccttg tgaatcctt tgaatgcaca caggatcct tgaatgcaca caggatcct tgaatgcaca tcataccttg ctataccttg ctataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg tgaatgcaca tcataccttg	tatatcacag ctgtaggggt ccaataggcc ttttttttt atcttagccc gctgatactg gcggggtatg tcgataagct gcataggct gcataggct tacttggagc acccgttct tacttggagc tcgccgacat tcgcggacat tcgcgcacat tcagccgttc ttatcatgca gctgcacat tcagccgtcat tcagctcac ccgcctcag cctcgctca tcagcagccat tcagcagcac tcagcagcac tcagcagcac tcagcagcac tcagcagcac ccgcacagcac cagcagcac cagcagcac cagcagcac	aggcatctat atcctgtttt ttctcctttt ttgtcacggc gcctctctgc ccgcgtttta ttaatgcggt ttaatgcgg ttaatgccg tcactatggc cactatcgac catcgtggcc catcgtggcc catcgtggca tatggtggca ccttgcgcataag gcttgcgtagga gcgtaggcg actcgtagga gcgtaggcg actcgtcgtc catggcgcc catggcgca agccttcgtc catggcggcc catggcgacg agccttcgtc catggcggcc cttccccatt gctgtcaag cgcgttgcgt taccagccta gagcacatgg cacagcgca aagcgcaaa acgcggcgca	7740 7800 7860 7920 7980 8040 8100 8160 8220 8340 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9120 9180 9360 9420 9480

gaacgtgaag	cgactgctgc	tgcaaaacgt	ctgcgacctg	agcaacaaca	tgaatggtct	9660
tcaatttcca	totttcgtaa	agtctggaaa	cgcggaagtc	agcgccctgc	accattatgt	9720
tccggatctg	catcgcagga	tgctgctggc	taccctgtgg	aacacctaca	tctgtattaa	9780
caaaacacta	gcattgaccc	tgagtgattt	ttctctggtc	ccgccgcatc	cataccgcca	9840
attatttacc	ctcacaacqt	tccaqtaacc	gggcatgttc	atcatcagta	acccgtatcg	9900
tgaggatcct	ctctcqtttc	atcggtatca	ttacccccat	gaacagaaat	tcccccttac	9960
acqqaqqcat	caaqtqacca	aacaqqaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
aagccagaca	ttaacqcttc	tggagaaact	caacgagctg	gacgcggatg	aacaggcaga	10080
catctgtgaa	tcacttcaca	accacqctga	tgagctttac	cgcagctgcc	tcgcgcgttt	10140
cootgatgac	ggtgaaaacc	tctgacacat	gcageteceg	gagacggtca	cagcttgtct	10200
gtaaggggat	accadaadca	gacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	10260
tcaaaacaca	gccatgaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	10320
gcggcatcag	agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tacataaqqa	gaaaataccg	catcaggcgc	tcttccgctt	cctcgctcac	tgactcgctg	10440
cactcaatca	ttcaactaca	gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat	caggggataa	cqcaqqaaaq	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaqqccqc	attactages	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcaacactc	aaqtcaqaqq	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	ccctqqaaq	ctccctcqtq	cgctctcctg	ttccgaccct	gccgcttacc	10740
ggatacctgt	ccacctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	10800
aggtatetea	attcaatata	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	10860
attcaaccca	accactacac	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	10920
cacqacttat	cgccactggc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	10980
aacaatacta	cagagttctt	gaagtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttggtatct	acactctact	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	11100
tccggcaaac	aaaccaccqc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
cocagaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	11220
tggaacgaaa	actcacqtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
tagatccttt	tctagataat	acgactcact	ata			11313
J	-	-				•

<210> 4 <211> 11313 <212> DNA <213> Artificial Sequence

<220> <223> Plasmid

#### <400> 4

60 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg tottcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180 gacgaccggg teetttettg gatcaacccg etcaatgeet ggagatttgg gegtgeecce 240 300 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 360 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 420 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480 540 ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600 660 cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780 840 cattegacea ccaagegaaa categeateg agegageaeg taeteggatg gaageeggte ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900 ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 960 gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1020 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080 ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140 1200 agegeatege ettetatege ettettgaeg agttettetg agtttaaaca gaecacaaeg gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260 cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1320 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct 1380 1440 aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1500 gttcctctgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg

aaccccccac	ctaacaacaa	atacctctac	ggccaaaaqc	cacgtgtata	agatacacct	1560
acsssaacaa	cacaacccca	gtgccacgtt	ntgagttgga	tagttgtgga	aaqaqtcaaa	1620
teatatat	gaagggtatt	2-200000000	gtgaaggatg	cccagaaggt	accccattot	1680
Lggurur i - h	caagegeace	taataasaa	testttast	atatttaata	agaattaaaa	1740
atgggatetg	accegggee	coggracaca	Lycettacat	gtgtttagtc	gaggeedada	1800
aacgtctagg	ccccccgaac	cacggggacg	Eggetteeet	ttgaaaaaca	cyacaacacc	
atggcgccta	ttacggccta	ctcccaacag	acgcgaggcc	tacttggctg	carcarcact	1860
agcctcacag	gccgggacag	gaaccaggtc	gagggggagg	tccaagtggt	ctccaccgca	1920
acacaatctt	tcctggcgac	ctgcgtcaat	ggcgtgtgtt	ggactgtcta	tcatggtgcc	1980
ggctcaaaga	cccttaccaa	cccaaaqqqc	ccaatcaccc	aaatgtacac	caatgtggac	2040
caggacctcg	teggetggea	agcgccccc	aaaacacatt	ccttgacacc	atgcacctgc	2100
addagaaaaa	acctttactt	aatracaaaa	cataccaata	tcattccggt	acaccaacaa	2160
ggcagcgcgg	aaaaaaaacat	actetecee	aggregatit	cctacttgaa	gaactettea	2220
ggcgacagca	tratetaca	ataggggg	aggeoggee	tettteagae	taccatatac	2280
ggeggteeae	Ligerergeee	ereggggeae	gergraggea	tctttcgggc	andadtata	2340
acccgagggg	ttgcgaaggc	ggrggaerri	gracecegreg	agtctatggg	adccaccacg	2400
cggtccccgg	tcttcacgga	caactcgtcc	cereeggeeg	taccgcagac	accodaggig	
gcccatctac	acgcccctac	tggtagcggc	aagagcacta	aggtgccggc	tgcgtatgca	2460
gcccaagggt	ataaggtgct	tgtcctgaac	ccgtccgtcg	ccgccaccct	aggtttcggg	2520
gcgtatatgt	ctaaggcaca	tggtatcgac	cctaacatca	gaaccggggt	aaggaccatc	2580
accacaaata	ccccatcac	qtactccacc	tatggcaagt	ttcttgccga	cggtggttgc	2640
tetagagaca	cctatgacat	cataatatqt	gatgagtgcc	actcaactga	ctcgaccact	2700
atecterera	tragracact	cctggacgaa	acadagacaa	ctggagcgcg	actcatcata	2760
atacagggaa	ctaccctcc	addatcaatc	accorporac	atccaaacat	caaaaaaata	2820
catatatata	ccacgcccc	aatacaattt	tatoocaaao	ccatccccat	cgagaggatc	2880
getetgteea	gcaccygaga	tttstssst	tacggcaaag	aatataataa	actractaca	2940
aagggggga	ggeaceteat	LLLCLGCCal	cccaagaaga	aatgtgatga	tatataaata	3000
aagctgtccg	gcctcggact	caargergra	geatattace	ggggccttga	cycattege	3060
ataccaacta	gcggagacgt	cattgtcgta	gcaacggacg	ctctaatgac	gggetttace	
ggcgatttcg	actcagtgat	cgactgcaat	acatgtgtca	cccagacagt	cgacttcagc	3120
ctggacccga	ccttcaccat	tgagacgacg	accgtgccac	aagacgcggt	gtcacgctcg	3180
cagcggcgag	gcaggactgg	taggggcagg	atgggcattt	acaggtttgt	gactccagga	3240
gaacggccct	cgggcatgtt	cgattcctcg	gttctgtgcg	agtgctatga	cgcgggctgt	3300
acttaataca	ageteacgee	cqccqaqacc	tcagttaggt	tgcgggctta	cctaaacaca	3360
ccadattac	ccatctacca	agaccaticta	gagttctggg	agagcgtctt	tacaggcctc	3420
addagatag	accccattt	cttatcccaa	actaaggagg	caggagacaa	cttcccctac	3480
acccacacag	acgeceatee	aatatacaca	addactcadd	ctccacctcc	atcotogoac	3540
ciggiageac	accaggerac	agagataaa	agggacaagg	acgggccaac	accetacta	3600
caaacgcgga	agracectar	acggccaaag	actacgeege	acgggccaac	casatacatc	3660
tataggetgg	gageegreea	aaacyayytt	attactacac	accccataac	catacacaca	3720
atggcatgca	tgtcagctga	ectggaggte	gccacgagca	cctgggtgct	ggtaggtgga	3780
gtcctagcag	ctctggccgc	gtattgcctg	acaacaggca	gcgtggtcat	Lytyyytayy	3840
atcatcttgt	ccggaaagcc	ggccatcatt	cccgacaggg	aagtctttta	ccgggagttc	
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacggggaat	gcagctcgcc	3900
gaacatttca	laacagaaggc	aatcgggttg	ctgcaaacag	ccaccaagca	agcggaggct	3960
getgeteeeg	cggtggaatc	caagtggcgg	accctcgaag	ccttctgggc	gaagcatatg	4020
tqqaatttca	tcagcgggat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
gcgatagcat	cactgatggc	attcacagcc	tctatcacca	gcccgctcac	cacccaacat	4140
accetectet	ttaacatcct	aaaaaaataa	ataaccaccc	aacttgctcc	tcccagcgct	4200
acttatactt	trataggege	caacatcact	agaacaacta	ttggcagcat	aggeettggg	4260
geeeegeee	togatattt	aggaagttat	aasacsaaaa	tggcaggcgc	actcataacc	4320
tttaaggtgataa	tasacaacas	gatacctcc	accdaddacc	tggttaacct	actccctqct	4380
-tt-t-	chagoggoga	actactaca	atcatataca	cagcgatact	acatraacac	4440
acceteteee	ceggegeeee	agregreggg	atanagaga	tastsacatt	cacttcacaa	4500
grgggeceag	gggagggggc	Lycagigg	acgaacegge	tgatagcgtt	agetetgegg	4560
ggtaaccacg	totoccocac	geactatgtg	cctgagageg	acgctgcagc	atgugucace	4620
cagatectet	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gragardaad	
gaggactgct	ccacgccatg	ctccggctcg	tggctaagag	atgtttggga	ttggatatgc	4680
acggtgttga	ctgatttcaa	gacctggctc	cagtccaagc	tcctgccgcg	attgccggga	4740
gtccccttct	tctcatgtca	acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	4800
caaaccacct	gcccatgtgg	agcacagatc	accggacatg	tgaaaaacgg	ttccatgagg	4860
atcataaaac	ctaggacctg	tagtaacacg	tggcatggaa	cattccccat	taacgcgtac	4920
accacadacc	cctacacacc	ctccccaaca	ccaaattatt	ctagggcgct	gtggcgggtg	4980
actactasaa	agtacgtgga	gattacacaa	atagagaatt	tccactacgt	gacgggcatg	5040
accactoaca	acatassata	cccatatcac	attecaacce	ccgaattctt	cacagaaqtq	5100
astacactas	acttaceas	atacactcas	acatacasa	ccctcctacg	agaagaaata	5160
garggggrgc	tagaaataaa	trastarre	attractasa	agctcccatg	Gaadacaaa	5220
acattectgg	ccgggcccaa	ttacatactcg	accesses a	ageneeattea	addadadada	5280
ccggacgtag	cagrgereae		accyaccect	cccacattac	addtatecad	5340
gctaagcgta	ggctggccag	gggatetece	- deceted tegg	ccagctcatc	agotatocag	5400
crgrcrgcgc	cttccttgaa	ggcaacatgc	actacccgtc	atgactcccc	ggacgccgac	
ctcatcgagg	ccaacctcct	gtggcggcag	gagatgggcg	ggaacatcac	ccgcgtggag	5460

tcagaaaata	aggtagtaat	tttqqaqtct	ttcgagccgc	tccaagcgga	ggaggatgag	5520
recagaaaaaa	ccgttccggc	adatecta	cadadataca	ggaaattccc	tcgagcgatg	5580
agggaagtat	ccgccccggc	-t-caracety	atattaa	agtectoraa	adsacadsa	5640
cccatatggg	cacgcccgga	ttacaaccct	ccactyttag	agicciggaa	ggacccggac	
tacgtccctc	cagtggtaca	cgggtgtcca	ttgccgcctg	ccaaggcccc	teegataeea	5700
cctccacqqa	qqaaqaqqac	ggttgtcctg	tcagaatcta	ccgtgtcttc	tgccttggcg	5760
gaggtrocca	caaagacctt	caacaactcc	gaatcgtcgg	ccgtcgacag	cggcacggca	5820
gagorogooa	ctgaccagcc	ctcccaccac	addascacaa	gatecgacgt	tgagtcgtac	5880
acggeerere	Clyactagec	ccccgacgac	ggcgacgcg	Statasaaas	agagtattaa	5940
tcctccatgc	cccccttga	gggggagccg	ggggarcecy	accidagega	egggeeeegg	
tctaccgtaa	gcgaggaggc	tagtgaggac	gtcgtctgct	gctcgatgtc	ctacacatgg	6000
acaddcdccc	tgatcacgcc	atgcgctgcg	gaggaaacca	agctgcccat	caatgcactg	6060
accaactctt	tgctccgtca	ccacaactto	gtctatgcta	caacatctcq	caqcqcaaqc	6120
agcaaccccc	agaaggtcac	atttananan	atacaaatac	tagacgacca	ctaccoccac	6180
ctgcggcaga	agaaggicac	cicigacaga	ctgtaggtee	changeter	ptacatagaa	6240
gtgctcaagg	agatgaaggc	gaaggcgtcc	acagttaagg	ClaadCLLCL	accogragag	
gaagcctgta	agetgaegee	cccacattcg	gccagatcta	aatttggcta	tggggcaaag	6300
gacgtccgga	acctatccag	caaggccgtt	aaccacatcc	gctccgtgtg	gaaggacttg	6360
aradonada	ctgagacacc	aattgacacc	accatcatgg	caaaaaatga	gattttctgc	6420
Ciggaagaca	ccgagacacc	aaccgacacc	cotaccetta	tcctattccc	agatttgggg	6480
grccaaccag	agaagggggg	degeaageea	guedgueda	cogcactece	tanagagata	6540
gttcgtgtgt	gcgagaaaat	ggccctttac	gatgtggtct	ecacectede	ccaggecgrg	
atgggctctt	catacggatt	ccaatactct	cctggacagc	gggtcgagtt	cctggtgaat	6600
acctagaaag	cgaagaaatg	ccctatgggc	ttcqcatatg	acacccgctg	ttttgactca	6660
Secasionary	agaatgacat	ccatattasa	gagtcaatct	accaatatta	tgacttggcc	6720
acggccactg	agaacgacac	cogegeeges	Sascara	tttacatcoo	agaccccta	6780
cccgaagcca	gacaggccat	aaggregere	acagagegge		tataataaaa	6840
actaattcta	aagggcagaa	ctgcggctat	cgccggtgcc	gegegagegg	Lycactyacy	
accadctdcd	gtaataccct	cacatgttac	ttgaaggccg	ctgcggcctg	tcgagctgcg	6900
aagctccagg	actgcacgat	actcatatac	ggagacgacc	ttgtcgttat	ctgtgaaagc	6960
aagaaaaaa	aagaggacga	addaadda	coogcettea	cagaaactat	gactagatac	7020
gegggaeee	aagaggacga	ggcgagccca	~	tagaattaat	aacatcatoo	7080
tatgacacaa	ctggggaccc	geccaaacca	gaacacgacc	t-t-sets	tatacaaagt	7140
tcctccaatg	tgtcagtcgc	gcacgatgca	tctggcaaaa	gggtgtacta	Leceaccege	
gaccccacca	cccccttgc	gcgggctgcg	tgggagacag	ctagacacac	tccagtcaat	7200
tectaactaa	gcaacatcat	catqtatqcq	cccaccttgt	gggcaaggat	gatcctgatg	7260
2000330003	tctccatcct	tctactcac	gaacaacttg	aaaaagccct	agattgtcag	7320
acctatttt		anttanaga	attanatta	ctcacatcat	traarractr	7380
atctacgggg	cctgttactc	carryaycca	telgatetae	ctcagaccac	teaactaat	7440
catggcctta	gcgcattttc	actccatagt	tactctccag	gtgagateaa	Lagggragger	
tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	7500
atccacacta	ggctactgtc	ccaqqqqqqq	agggctgcca	cttgtggcaa	gtacctcttc	7560
aactaaacaa	taaggaccaa	gctcaaactc	actccaatcc	caactacatc	ccaqttqqat	7620
aactgggcag	ggttcgttgc	taattaaaa	aaaaaaaaaa	tatatcacad	cctatctcat	7680
ttateeaget	ggttegttge	Lygicacage	9999949464	atatacaaaa	aggatatat	7740
gcccgacccc	gctggttcat	grggrgccra	CLCCLACLLL	Crgragggr	aggcacctac	
ctactcccca	accgatgaac	ggggacctaa	acactccagg	ccaataggcc	accetgeee	7800
tttccctttt	tttttttttt	tttttttt	ttttttttt	tttttttt	ttctccttt	7860
ttttcctct	ttttttcctt	ttctttcctt	taataactcc	atcttagccc	tagtcacggc	7920
tocctcccc	aggtccgtga	accacttasc	tacagagagt	actaatacta	acctctctac	7980
Laguigigaa	aggiccgiga	geegeeegae	cgcagagaga	googaatata	ccccctttta	8040
agatcaagta	ctcctgcagg	egegecacia	grgggaarac	geggggeacg	thickere	8100
gcatattgac	gacccaattc	tcatgtttga	cagcttatca	tegataaget	LLaatgeggt	
agtttatcac	agttaaattg	ctaacgcagt	caggcaccgt	gtatgaaatc	taacaatgcg	8160
ctcatcgtca	tcctcggcac	catcacccta	gatgctgtag	gcataggctt	ggttatgccg	8220
ataataaaa	gcctcttgcg	ggatatcotc	cattcccaca	acatcaccaa	tcactatggc	8280
gracigorgg	cgctatatgc	attantago	tttatataca	cacccattct	cogaggacto	8340
grgcrgcrag	egetatatge	gregargeaa	-tt	tactogecoe	anathtana	8400
tccgaccgct	ttggccgccg	cccagrecty	etegettege	Lacitygage	caccaccgac	8460
tacgcgatca	tggcgaccac	acccgtcctg	tggatcctct	acgccggacg	categragee	
gacaticacca	acaccacaaa	tacaattact	ggcgcctata	tcgccgacat	caccgatggg	8520
assastaaa	ctcgccactt	cagactcata	agcacttatt	tcaacataaa	tatqqtqqca	8580
gaagaccggg	ccgggggact	appacaaa	atataattaa	atocaccatt	cettacaaca	8640
ggeeeegrag	ccgggggact	gulgggegee	acceceege	tastassass	ataaataaa	8700
gcggtgctca	acggcctcaa	cctactactg	ggetgettee	Laatytayga	geogeacaag	
ggagagcgtc	gaccgatgcc	cttgagagcc	ttcaacccag	tcagctcctt	ccggrgggcg	8760
cagaacataa	ctatcgtcgc	cgcacttatg	actgtcttct	ttatcatgca	actcgtagga	8820
cadataccad	cagcgctctg	ggtcattttc	gacgaggacc	actttcgctg	gagcgcgacg	8880
-~35~5~535	tgtcgcttgc	agtattegga	atchtgrace	ccctcactca	agcettegte	8940
acyaccyycc	-gaugerege	13 -ac-cyga	222222222	ttatagaaaa	cataacaaca	9000
actggtcccg	ccaccaaacg	rrrayyayay	aagcaggcca		attacaatt	9060
gacgcgctgg	gctacgtctt	getggegtte	gcgacgcgag	gctggatggc	GLLCCCCart	
atgattcttc	tegetteegg	cggcatcggg	atgcccgcgt	tgcaggccat	gctgtccagg	9120
caggtagatg	acgaccatca	gggacagctt	caaqqatcqc	tegeggetet	taccagccta	9180
2015003cc3	ctdcacccc+	gatcgtcacg	gcgatttatg	cagaataaaa	gagcacatgg	9240
accurgated	. coggactyct	244424444	atatacette	+0+000+000	cacattacat	9300
aacgggttgg	catggattgt	aggegeegee			ogogogogo	9360
cgcggtgcat	ggagccgggc	cacctcgacc	cgaacggaag	ceggeggeae	ctcgctaacg	
gattcaccac	tccaagaatt	ggagccaatc	aattcttgcg	gagaactgtg	aatgcgcaaa	9420
-	-					

ccaacccttq	gcagaacata	tccatcqcqt	ccgccatctc	cagcagccgc	acgcggcgca	9480
tetegggeag	cattagatee	tggccacqqq	tgcgcatgat	cgtgctcctg	tcgttgagga	9540
cccaactaga	ctaacaaaat	tgccttactq	qttagcagaa	tgaatcaccg	atacgcgagc	9600
gaacgtgaag	cgactgctgc	tgcaaaacgt	ctgcgacctg	agcaacaaca	tgaatggtct	9660
tcgatttccg	tgtttcgtaa	agtctggaaa	cgcggaagtc	agcgccctgc	accattatgt	9720
tccqqatctq	catcgcagga	tgctgctggc	taccctgtgg	aacacctaca	tctgtattaa	9780
cgaagcgctg	gcattgaccc	tgagtgattt	ttctctggtc	ccgccgcatc	cataccgcca	9840
attatttacc	ctcacaacgt	tccagtaacc	gggcatgttc	atcatcagta	acccgtatcg	9900
tgagcatcct	ctctcgtttc	atcggtatca	ttacccccat	gaacagaaat	tcccccttac	9960
acqqaqqcat	caagtgacca	aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
aaqccaqaca	ttaacqcttc	tggagaaact	caacgagctg	gacgcggatg	aacaggcaga	10080
catctqtqaa	tcgcttcacg	accacgctga	tgagctttac	cgcagctgcc	tegegegttt	10140
cagtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	cagcttgtct	10200
qtaaqcqqat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	10260
tcqqqqcqca	qccatgaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	10320
acaacatcaa	agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tacataaaaa	gaaaataccg	catcaggcgc	tcttccgctt	cctcgctcac	tgactcgctg	10440
cactcaatca	ttcggctgcg	gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat	caqqqqataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	10740
ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	10800
aggtatctca	gttcggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	10860
gttcagcccg	accgctgcgc	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	10920
cacqacttat	cgccactggc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	10980
ggcggtgcta	cagagttctt	gaagtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttggtatct	acactctact	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	11100
tccqqcaaac	aaaccaccgc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
cacagaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	11220
tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
tagatccttt	tctagataat	acgactcact	ata		•	11313

<210> 5 <211> 11313 <212> DNA

<213> Artificial Sequence

<220> <223> Plasmid

### <400> 5

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120 180 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 240 gacgaccggg teetttettg gatcaacccg etcaatgeet ggagatttgg gegtgeecee gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360 420 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480 ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540 acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600 cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660 tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780 840 cattegacea ceaagegaaa categeateg agegageaeg taeteggatg gaageeggte 900 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 960 gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1020 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080 ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140 1200 agegeatege ettetatege ettettgaeg agttettetg agtttaaaca gaecacaaeg 1260 qtttccctct agcqqqatca attccgcccc tctccctccc cccccctaa cgttactggc cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1320

						4 7 7 7
ccgtcttttg	gcaatgtgag	ggcccggaaa	cctggccctg	tcttcttgac	gagcattcct	1380
agggtgttt	aggetetece	casaccasta	caaggtetat	tgaatgtcgt	gaaggaagga	1440
aggggcccc		ouduggaaog				1500
gttcctctgg	aagcttcttg	aagacaaaca	acgreegeag	cgaccctttg	caggcagcgg	
aaccccccac	ctggcgacag	atacctctac	ggccaaaagc	cacgtgtata	agatacacct	1560
	22-2	atacascatt	atasattaas	tagttgtgga	aarartraaa	1620
gcaaaggcgg	Cacaacccca	gegecacgee	grageraga	tagttgtgga	aagagccaaa	
tageteteet	caagcgtatt	caacaagggg	ctgaaggatg	cccagaaggt	accccattgt	1680
ataggateta	atetagaace	trantaraca	toctttacat	gtgtttagtc	gaggttaaaa	1740
acgggacceg	acceggggee	ccggcgcaca		5-5	5055	1800
aacgtctagg	ccccccgaac	cacggggacg	Eggtttteet	ttgaaaaaca	egataatace	
atogogota	ttacqqccta	ctcccaacag	acqcqaqqcc	tacttggctg	catcatcact	1860
409909000	~~~~~~~~	annagaaata	asaaaaaaaa	transatant	ataaaaaaa	1920
ageeteacag	geegggaeag	gaaccaggte	909999999	tccaagtggt	cccaccgca	
acacaatctt	tcctggcgac	ctgcgtcaat	ggcgtgtgtt	ggactgtcta	tcatggtgcc	1980
aratassas	carttaccaa	cccasadaac	ccaatcaccc	aaatgtacac	caatgtggac	2040
ggcccaaaga	ccccage	cccaaagggo				2100
caggacctcg	tcggctggca	agegeeeece	ggggcgcgtt	ccttgacacc	argeacerge	
agcaget.cgg	acctttactt	gatcacgagg	catqccqatq	tcattccggt	gcgccggcgg	2160
550050000	aaaaaaaa	atatagaaa	aggregatet	cctacttcaa	anactettea	2220
ggcgacagca	gggggagccc	acticititi	aggcccgccc	cctacttgaa	99900000	
ggcggtccac	tgctctgccc	ctcggggcac	gctgtgggca	tettteggge	tgccgtgtgc	2280
addedadddd	ttacasaaac	aataaacttt	atacccatca	agtctatggg	aaccactato	2340
accegaggg	ccgcgaagge	990994066	500000500	490000000	attacacata	2400
cggtccccgg	tcttcacgga	caactcgtcc	cctccggccg	taccgcagac	accodaggig	
gcccatctac	acocccctac	taataacaac	aagagcacta	aggtgccggc	tgcgtatgca	2460
		tataatassa	agatagatag	acaccaccct	aggtttgggg	2520
gcccaagggt	acaaggigei	Lycolyaac	cogcocycog	ccgccaccct	aggereeggg	
gcgtatatgt	ctaaggcaca	tggtatcgac	cctaacatca	gaaccggggt	aaggaccatc	2580
accaccacta	ccccatcac	gtactccacc	tatogcaagt	ttcttgccga	caataattac	2640
accacgggcg	CCCCaccac	generate			-55-55-5-	2700
tctgggggcg	cctatgacat	cataatatgt	gatgagtgcc	actcaactga	Cityaccaci	
atcctgggca	tcggcacagt	cctggaccaa	gcggagacgg	ctggagcgcg	actcgtcgtg	2760
atacacacac	ataaaaataa	aggetcagtc	accoraccac	atccaaacat	caaaaaaata	2820
cicyclaccy	Clacycoloc	9994669966	h-b		-3-33-33-3	2880
gctctgtcca	gcactggaga	aacccccttt	catggcaaag	ccatccccat	cyayaccacc	
aaggggggga	ggcacctcat	tttctqccat	tccaagaaga	aatgtgatga	gatagaagag	2940
anaatataaa	acatacaact	caatoctota	gcatattacc	ggggccttga	totatecote	3000
aagetgteeg	geereggaer.	caacgoogca	500000000	333344434	acastttaaa	3060
ataccaacta	gcggagacgt	cattgtcgta	geaacggacg	ctctaatgac	gggcttacc	
gacgatttcg	actcaqtqat	cgactgcaat	acatgtgtca	cccagacagt	cgacttcagc	3120
ataraaraa	cetteaceat	taaaacaaca	accotoccac	aagacgcggt	atcacactca	3180
ctggattega		cgagacgacg	-1		Jeer-Jeer J	3240
cagcggcgag	gcaggactgg	taggggcagg	atgggcattt	acaggtttgt	gactecagga	
gaacggccct	cqqqcatqtt	cgattcctcg	gttctgtgcg	agtgctatga	cgcgggctgt	3300
acttaataca	aggtgaggg	caccasasca	tcagttaggt	tgcgggctta	cctaaacaca	3360
geeeggeacg	ugoccuogoo	0300303000	~~attataa	agaggtgtt	tacacccccc	3420
ccagggttgc	cegtetgeea	ggaccatctg	gagereeggg	agagcgtctt		
acccacatag	acgcccattt	cttgtcccag	actaagcagg	caggagacaa	cttcccctac	3480
ctggtagcat	accaggetae	aatatacacc	agggctcagg	ctccacctcc	atcgtgggac	3540
annatataan	actatatast	acaactaaaa	cctacactac	acgggccaac	accectacta	3600
caaacycyga	agigicicat	acggccaaag	cccacgccgc	2099900000	goodagata	3660
tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accccataac	Caaalacatt	
atggcatgca	tatcaactaa	cctqqaqqtc	qtcacgagca	cctgggtgct	ggtaggcgga	3720
ataataaaa	atataaaaaa	atattaccta	acaacacca	gcgtggtcat	tataaacaaa	3780
greerageag	ccccggccgc	geacegeeeg	acaacaggea	3030330000	-5-555-55	3840
atcatcttgt	ccggaaagcc	ggccatcatt	eccgacaggg	aagtctttta	ccgggagttc	
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacggggaat	gcagctcgcc	3900
gaagatttga	22222	aatcaaatta	ctocaaacao	ccaccaagca	agcagaaact	3960
gaacacccca	aacagaaggc	aaccgggccg	cegouadoug		~5~555	4020
gctgctcccg	cggtggaatc	caagtggcgg	accetegaag	ccttctgggc	yaaycacacy	
togaatttca	tcaqcqqqat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
aggataggat	aactaataac	attrarancr	totatoacca	gcccgctcac	cacccaacat	4140
gcgacagcac	caccgacggc	accoacagoo	~~~~~~~	onattaataa	tagasagaat	4200
accetectgt	ttaacatcct	ggggggargg	gradecacec	aacttgctcc	LCCCageger	
acttctactt	tcqtaqqcqc	cggcatcgct	ggagcggctg	ttggcagcat	aggccttggg	4260
anataatta	tagatatttt	agcaggttat	adadcadada	tggcaggcgc	actcataacc	4320
aagguguuu	tygututtet	9900990000	33~3~3333	+~~++>>~~+	ant agatagt	4380
tttaaggtca	taaacaacaa	gatgeeetee	accuauuacc	Luullaauu	acticities	
	-3-3-3-3-			- 3		
atcctctccc	ctggcgccct	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4440
atcetetece	ctggcgccct	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4500
gtgggcccag	ctggcgccct gggaggggc	agtcgtcggg tgtgcagtgg	gtcgtgtgcg atgaaccggc	cagcgatact tgatagcgtt	gcgtcggcac cgcttcgcgg	4500
gtgggcccag ggtaaccacg	ctggegeeet gggagggge teteeecae	agtcgtcggg tgtgcagtgg gcactatgtg	gtcgtgtgcg atgaaccggc cctgagagcg	cagcgatact tgatagcgtt acgctgcagc	gegteggeae egettegegg aegtgteaet	4500 4560
gtgggcccag ggtaaccacg	ctggegeeet gggagggge teteeecae	agtcgtcggg tgtgcagtgg gcactatgtg	gtcgtgtgcg atgaaccggc cctgagagcg	cagcgatact tgatagcgtt acgctgcagc	gegteggeae egettegegg aegtgteaet	4500 4560 4620
gtgggcccag ggtaaccacg cagatcctct	ctggcgcct gggaggggc tctccccac ctagtcttac	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga	cagcgatact tgatagcgtt acgctgcagc ggcttcacca	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac	4500 4560
gtgggcccag ggtaaccacg cagatcctct gaggactgct	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc	4500 4560 4620 4680
gtgggccag ggtaaccacg cagatcctct gaggactgct acggtgttga	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga	4500 4560 4620 4680 4740
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg	4500 4560 4620 4680 4740 4800
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg	4500 4560 4620 4680 4740
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg	4500 4560 4620 4680 4740 4800 4860
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct atcgtggggc	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc tagtaacacg	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg tggcatggaa	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg cattccccat	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg taacgcgtac	4500 4560 4620 4680 4740 4800 4860 4920
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct atcgtggggc	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg tggcatggaa ccaaattatt	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg cattccccat ctagggcgct	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg taacgcgtac gtggcgggtg	4500 4560 4620 4680 4740 4800 4860 4920 4980
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct atcgtggggc	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg tggcatggaa ccaaattatt	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg cattccccat ctagggcgct	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg taacgcgtac gtggcgggtg	4500 4560 4620 4680 4740 4800 4860 4920
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct atcgtgggc accacgggcc gctgctgagg	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc agtacgtgga	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg ggttacgcgg	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg tggcatggaa ccaaattatt gtgggggatt	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg cattccccat ctagggcgct tccactacgt	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg taacgcgtac gtggcgggtg gacgggcatg	4500 4560 4620 4680 4740 4800 4920 4980 5040
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct atcgtgggc accacgggcc gctgctgagg accactgaca	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc agtacgtgga acgtaaagtg	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg ggttacgcgg cccgtgtcag	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg tggcatggaa ccaaattatt gtgggggatt gttccggcc	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg cattccccat ctagggcgct tccactacgt	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg taacgcgtac gtggcgggtg gacgggcatg cacagaagtg	4500 4560 4620 4680 4740 4860 4920 4980 5040 5100
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct atcgtgggc accacgggcc gctgctgagg accactgaca gatggggtgc	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc agtacgtgga acgtaaagtg	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg ggttacgcgg cccgtgtcag gtacgctca	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg tggcatggaa ccaaattatt gtgggggatt gttccggccc gcgtgcaaac	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg cattccccat ctagggcgct tccactacgt ccgaattctt ccctcctacg	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg taacgcgtac gtggcgggtg gacgggcatg cacagaagtg ggaggaggtc	4500 4560 4620 4680 4740 4860 4920 4980 5040 5160
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct atcgtgggc accacgggcc gctgctgagg accactgaca gatggggtgc	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc agtacgtgga acgtaaagtg	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg ggttacgcgg cccgtgtcag gtacgctca	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg tggcatggaa ccaaattatt gtgggggatt gttccggccc gcgtgcaaac	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg cattccccat ctagggcgct tccactacgt ccgaattctt ccctcctacg	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg taacgcgtac gtggcgggtg gacgggcatg cacagaagtg ggaggaggtc	4500 4560 4620 4680 4740 4860 4920 4980 5040 5160 5220
gtgggcccag ggtaaccacg cagatcctct gaggactgct acggtgttga gtccccttct caaaccacct atcgtgggc accacgggcc gctgctgagg accactgaca gatggggtgc acattcctgg	ctggcgcct gggaggggc tctccccac ctagtcttac ccacgccatg ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc agtacgtgga acgtaaagtg ggttgcacag tcgggctcaa	agtcgtcggg tgtgcagtgg gcactatgtg catcactcag ctccggctcg gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg ggttacgcgg cccgtgtcag gtacgctcca tcaatacctg	gtcgtgtgcg atgaaccggc cctgagagcg ctgctgaaga tggctaagag cagtccaagc aagggagtct accggacatg tggcatggaa ccaaattatt gtgggggatt gttccggccc gcgtgcaaac gttgggtcac	cagcgatact tgatagcgtt acgctgcagc ggcttcacca atgtttggga tcctgccgcg ggcggggcga tgaaaaacgg cattccccat ctagggcgct tccactacgt	gcgtcggcac cgcttcgcgg acgtgtcact gtggatcaac ttggatatgc attgccggga cggcatcatg ttccatgagg taacgcgtac gtggcgggtg gacgggcatg cacagaagtg ggaggaggtc cgagcccgaa	4500 4560 4620 4680 4740 4860 4920 4980 5040 5160

gctaagcgta	ggctggccag	gggatetece	ccctccttgg	ccagctcatc	agctatccag	5340
	cttccttgaa					5400
	ccaacctcct					5460
tcagaaaata	aggtagtaat	tttggagtct	ttcgagccgc	tccaaqcqqa	agaggatgag	5520
						5580
	ccgttccggc					
cccatatggg	cacgcccgga	ttacaaccct	ccactgttag	agtcctggaa	ggacccggac	5640
tacqtccctc	cagtggtaca	caaatataca	ttaccaccta	ccaaggcccc	teccatacca	5700
	ggaagaggac					5760
qaqctcqcca	caaagacctt	caacaactcc	gaatcgtcgg	ccqtcqacaq	caacacaaca	5820
	ctgaccagcc					5880
tcctccatgc	cccccttga	gggggagccg	ggggatcccg	atctcagcga	cgggtcttgg	5940
	gcgaggaggc					6000
	tgatcacgcc					6060
agcaactctt	tgctccgtca	ccacaacttq	gtctatgcta	caacatctcq	cagcgcaagc	6120
	agaaggtcac					6180
gtgctcaagg	agatgaaggc	gaaggcgtcc	acagttaagg	ctaaacttct	atccgtggag	6240
gaageetgta	agctgacgcc	cccacattca	gccagatcta	aatttggcta	tagagaaaaa	6300
	acctatccag					6360
ctggaagaca	ctgagacacc	aattgacacc	accatcatgg	caaaaaatga	ggttttctgc	6420
	agaagggggg					6480
	gcgagaaaat					6540
atgggctctt	catacggatt	ccaatactct	cctggacagc	agatcaaatt	cctaataaat	6600
						6660
	cgaagaaatg					
acggtcactg	agaatgacat	ccgtgttgag	gagtcaatct	accaatgttg	tgacttggcc	6720
cccgaagcca	gacaggccat	aaggtcgctc	acadadcddc	tttacatcoo	agacccccta	6780
						6840
	aagggcagaa					
accagctgcg	gtaataccct	cacatgttac	ttgaaggccg	ctgcggcctg	tcgagctgcg	6900
	actgcacgat					6960
						7020
	aagaggacga					
tatgacaca	ctggggaccc	gcccaaacca	gaatacgact	tggagttgat	aacatcatgc	7080
tcctccaato	tgtcagtcgc	acacaataca	tctqqcaaaa	agatatacta	tctcacccqt	7140
	cccccttgc					7200
	gcaacatcat					7260
actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaaagccct	agattgtcag	7320
	cctgttactc					7380
	gcgcattttc					7440
tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	7500
aticcacacta	ggctactgtc	ccaaaaaaaa	agggctgcca	cttgtgggaa	gtacctcttc	7560
	taaggaccaa					7620
ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	7680
	gctggttcat					7740
	accgatgaac					7800
tttccctttt	tttttttctt	ttttttttt	ttttttttt	ttttttttt	ttctcctttt	7860
tttttcctct	ttttttcctt	ttettteett	taataactcc	atcttagccc	tagtcacggc	7920
						7980
	aggtccgtga					
agatcaagta	ctcctgcagg	cgcgccacta	gtgggaatac	gcggggtatg	ccgcgtttta	8040
	gacccaattc					8100
						8160
	agttaaattg					
ctcatcgtca	tcctcggcac	cgtcaccctg	gatgctgtag	gcataggctt	ggttatgccg	8220
	gcctcttgcg					8280
	cgctatatgc					8340
tccgaccgct	ttggccgccg	cccagtcctg	ctcgcttcgc	tacttqqaqc	cactatcgac	8400
	tggcgaccac					8460
	gcgccacagg					8520
gaagatcggg	ctcgccactt	cgggctcatg	agcgcttgtt	tcggcgtggg	tatggtggca	8580
	ccgggggact					8640
22222222	555555			++		8700
gugguguta	acggcctcaa	culautautg	ggorgottoo	Laatycagga	gucycataag	
ggagagcgtc	gaccgatgcc	cttgagagcc	ttcaacccag	tcagctcctt	ccggtgggcg	8760
cagaacataa	ctatcgtcgc	cocacttato	actotetet	ttatcatoca	actcotagga	8820
- 22222222	daddad+a+~	cot cotte	~~~~~~~~			8880
cayyryccyy	cagcgctctg	ggicalitic	ggcgaggadd	gerregerg	gagugugaug	
atgatcggcc	tgtcgcttgc	ggtattcgga	atcttgcacg	ccctcgctca	agccttcgtc	8940
	ccaccaaacg					9000
usuuuuutua Du	aut saat at t	autacatta		adtactec	4+44444	9060
	gctacgtctt					
atgattcttc	tegetteegg	cggcatcggg	atgcccgcgt	tgcaggccat	gctgtccagg	9120
	acgaccatca					9180
						9240
uccegacoa	ctggaccgct	garrace	gogactiaty	cogocoogge	gagcacacgg	2640

aacgggttgg catggat	tgt aggcgccgcc	ctataccttg	tetgeeteee	cgcgttgcgt	9300
cgcggtgcat ggagccg	ggc cacctcgacc	tgaatggaag	ccggcggcac	ctcgctaacg	9360
gattcaccac tccaaga	att ggagccaatc	aattcttgcg	gagaactgtg	aatgcgcaaa	9420
ccaaccettg gcagaac	ata tccatcgcgt	ccgccatctc	cagcagccgc	acgcggcgca	9480
tctcgggcag cgttggg	tcc tggccacggg	tgcgcatgat	cgtgctcctg	tcgttgagga	9540
cccggctagg ctggcgg	ggt tgccttactg	gttagcagaa	tgaatcaccg	atacgcgagc	9600
gaacgtgaag cgactgc	tgc tgcaaaacgt	ctgcgacctg	agcaacaaca	tgaatggtct	9660
toggtttccg tgtttcg	taa agtctggaaa	cgcggaagtc	agcgccctgc	accattatgt	9720
teeggatetg categoa	gga tgctgctggc	taccctgtgg	aacacctaca	tctgtattaa	9780
cgaagcgctg gcattga	ccc tgagtgattt	ttctctggtc	ccgccgcatc	cataccgcca	9840
gttgtttacc ctcacaa	cgt tccagtaacc	gggcatgttc	atcatcagta	acccgtatcg	9900
tgagcatcct ctctcgt	ttc atcggtatca	ttacccccat	gaacagaaat	tcccccttac	9960
acggaggcat caagtga	cca aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
aagccagaca ttaacgc	ttc tggagaaact	caacgagctg	gacgcggatg	aacaggcaga	10080
catctgtgaa tcgcttc	acg accacgctga	tgagctttac	cgcagctgcc	tcgcgcgttt	10140
cagtaataac gataaa	acc tctgacacat	gcagctcccg	gagacggtca	cagcttgtct	10200
gtaagcggat gccggga	.qca qacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	10260
togggggga gccatga	.ccc agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	10320
gcggcatcag agcagat	tgt actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tocotaagga gaaaata	.ccq catcaggcgc	tetteegett	cctcgctcac	tgactcgctg	10440
cacteagtea ttegget	gcg gcgagcggta	. tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat cagggga	taa cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta aaaaggc	cgc gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa atcgacg	ctc aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	10680
caggogtttc cccctgg	aag ctccctcgtg	geteteetg	ttccgaccct	gccgcttacc	10740
ggatacctgt ccgcctt	tct cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	10800
aggtatetea gtteggt	gta ggtcgttcgc	: tccaagctgg	gctgtgtgca	cgaacccccc	10860
gttcagcccg accgctg	ıcgc cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	10920
cacgacttat cgccact	ggc agcagccact	. ggtaacagga	ttagcagagc	gaggtatgta	10980
ggcggtgcta cagagtt	ctt gaagtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttggtatct gcgctct	gct gaagccagtt	: accttcggaa	aaagagttgg	tagctcttga	11100
tccggcaaac aaaccac	cgc tggtagcggt	: ggtttttttg	tttgcaagca	gcagattacg	11160
cqcaqaaaaa aaqqatc	tca agaagatcct	: ttgatctttt	ctacggggtc	tgacgctcag	11220
tggaacgaaa actcacg	ıtta agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
tagateettt tetagat	aat acgactcact	: ata			11313

<210> 6 <211> 11313 <212> DNA <213> Artificial Sequence

<220> <223> Plasmid

### <400> б

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 tetteaegea gaaagegtet ageeatggeg ttagtatgag tgtegtgeag eeteeaggae 120 180 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag gacgaccggg teetttettg gatcaacccg etcaatgeet ggagatttgg gegtgeecee 240 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300 gtgettgega gtgeeeggg aggtetegta gaeegtgeae catgageaeg aateetaaae 360 420 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480 ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540 acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600 cgacgggcgt tecttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660 tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780 cattegacea ccaagegaaa categeateg agegageaeg tacteggatg gaageeggte 840 900 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 960 ccaggeteaa ggegegeatg cccgaeggeg aggatetegt egtgaeceat ggegatgeet 1020 gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080 ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140

adcacaticac	cttctatcgc	cttcttgacg	agttcttctg	agtttaaaca	gaccacaacg	1200
		***************************************	teteestage	adadadataa	aattaataaa	1260
gttteectet	agcgggatca	acceegeeee	LULUGULLUL	CCCCCCCaa	Cyccaccygc	
cgaagccgct	tggaataagg	ccggtgtgcg	tttgtctata	tgttattttc	caccatattg	1320
agatettta	gcaatgtgag	aacccaassa	cetaacceta	tettettgae	gaggattcct	1380
Cogcococo	goddogogdg	9900099444		tanatataat	20200000	1440
aggggtcttt	cccctctcgc	caaaggaatg	caaggtetgt	Lyaaryrcyr	gaaggaagca	
attectetaa	aagcttcttg	aaqacaaaca	acqtctqtaq	cgaccctttg	caqqcaqcqq	1500
30000000	ctggcgacag	ataaatataa	uuuusssauu	cacatatata	anatanannt	1560
aaccccccac	Ciggigadag	graceracia	ggccaaaagc	cacgigiaia	agacacaccc	
qcaaaggcgg	cacaacccca	gtgccacgtt	gtgagttgga	tagttgtgga	aagagtcaaa	1620
taggtgtggt	caagcgtatt	caacaadddd	chgaaggatg	cccagaaggt.	accccattqt	1680
Lygorotot	caagegeace	caacaagggg		~~~	~~~~	
atgggatctg	atctggggcc	teggtgeaca	tgetttacat	gigillayio	gaggitaaaa	1740
aacotctaoo	cccccgaac	cacqqqqacq	tggttttcct	ttgaaaaaca	cgataatacc	1800
	ttacggccta	ataganaan	addedadddd	tacttaaata	catcatcact	1860
atggegeeta	LLacygodla	CLCCCaacag	acgcgaggcc	Lacteggerg	·	
agcctcacag	gccgggacag	gaaccaggtc	gagggggagg	tccaagtggt	ctccaccgca	1920
acacaatett	tcctggcgac	ctacatcaat	agcatatatt.	ggactgtcta	tcatootocc	1980
acacaaccc	2000550540		22-2-2-2-	222+4+2424	asstataasa	2040
ggctcaaaga	cccttgccgg	cccaaagggc	CCaaccaccc	adatytatat	caacycygac	
caggacctcg	tcggctggca	agcqcccccc	ggggcgcgtt	ccttgacacc	atgcacctgc	2100
323	acctttactt	aataaaaaa	cataccaata	trattroggt	acaccaacaa	2160
ggcagcgcgg	acciticacti	ggccacgagg	caegocgaeg	coaccogge	3030033033	
ggcgacagca	gggggagcct	actctccccc	aggcccgttt	cctacttgaa	gggctcttcg	2220
aacaatccac	tgctctgccc	chagagacac	actataaaca	tctttcaaac	taccatatac	2280
9909900000	th		2-2-2-2-	actatataca		2340
acccgagggg	ttgcgaaggc	ggtggacttt	graceegreg	agicialygg	aaccaccacy	
caatccccaa	tcttcacgga	caactcgtcc	cctccggccg	taccgcagac	attccaggtg	2400
aggastatag	acgcccctac	tantanenae	aadadcacta	aggtaccage	tacatataca	2460
goodatotac	acgecectae	Lyguagegge	uagagoaota	4330300330		
gcccaagggt	ataaggtgct	tgtcctgaac	cegteegteg	ccgccaccct	aggttteggg	2520
acatatatat	ctaaggcaca	taatatcaac	cctaacatca	gaaccggggt	aaqqaccatc	2580
20200000	ccccatcac	atactacaca	tataaraaat	ttcttacca	caataattac	2640
accacgggcg	Coccarcac	gracicacc	catggeaage	ccccgccga	0990950050	
tctgggggcg	cctatgacat	cataatatgt	gatgagtgcc	actcaactga	etegaeeaet	2700
atcctgggca	tcggcacagt	cctggaccaa	gcggagacgg	ctggagcgcg	actcgtcgtg	2760
atararara	ctacgcctcc	aggategate	accotoccac	atccaaacat	caaaaaaata	2820
decycladeg	clacycoloc	gggaccggcc	accycyccac		0303303303	2880
gctctgtcca	gcactggaga	aatccccttt	tatggcaaag	CCalccccal	cyagaccarc	
aaqqqqqqa	ggcacctcat	tttctqccat	tccaagaaga	aatgtgatga	getegeegeg	2940
2222222	gcctcggact	castrotota	gratattacc	agaacettaa	totatecote	3000
aagergreeg	geeeeggaee	oud og o og o d	500000000	22222-	-JJ	3060
ataccaacta	gcggagacgt	cattgtegta	gcaacggacg	Ciclaalyac	gggctttacc	
ggcgatttcg	actcagtgat	cgactgcaat	acatgtgtca	cccagacagt	cgacttcagc	3120
ctagacccga	ccttcaccat	taaaacaaca	accotoccac	aagacgcggt	gtcacgctcg	3180
	gcaggactgg	-2-2-2-2	atagagattt	acacattat	dactccadda	3240
cageggegag	geaggaetgg	Laggggcagg	acgggcaccc	acaggeeege	9400004554	
gaacggccct	cgggcatgtt	cgattcctcg	gttctgtgcg	agtgctatga	egegggetgt	3300
acttaataca	agctcacgcc	cqccqaqacc	tcaqttaqqt	tgcgggctta	cctaaacaca	3360
ganagattag	ccgtctgcca	agaccateta	ganttetaga	agagggtgtt	tacaggcctc	3420
ccagggccgc	ccgcccgcca	ggaccacceg	3490000333		attaggetsa	3480
acccacatag	acgcccattt	cttgtcccag	accaagcagg	Cayyayacaa	CLLCCCCCac	
ctqqtaqcat	accaggctac	ggtgtgcgcc	agggctcagg	ctccacctcc	atcgtgggac	3540
casatataas	agtgtctcat	accortaaaa	cctacactac	acoooccaac	acccctacta	3600
cadacgegga	agagaaaaa			agggataag	gaaatagatg	3660
tataggetgg	gagccgttca	aaacgaggct	actaccacac	accicataac		
atggcatgca	tgtcagctga	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	3720
atcetageag	ctctggccgc	gtattgcctg	acaacaggca	acataatcat	tataaacaaa	3780
geoceagoag		5000050005		22244+++2	agaggagtta	3840
atcatcttgt	ccggaaagcc	ggccarcarr	cccgacaggg	adgittitta	ccgggagttc	
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacggggaat	gcagctcgcc	3900
gaacatttca	aacagaaggc	aatcqqqttq	ctocaaacao	ccaccaaqca	agcggaggct	3960
		annatagaaa	addatadaad	aattataaa	gaaggatatg	4020
getgeteeeg	cggtggaatc	caaguggugg	accecegaag	cccccgggc	gaagcacacg	
tggaatttca	tcagcgggat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
acastaacat	cactgatggc	attcacaccc	tctatcacca	acccactcac	cacccaacat	4140
5054045040	ttaacatcct	aaaaaastaa	ataaaaaaaa	aacttoctcc	teceageget	4200
accetectgt	LLdaCalCCL	ggggggacgg	graderace	aactigetee	ccccagegee	
gcttctgctt	tcgtaggcgc	cggcatcgct	ggagcggctg	ttggcagcat	aggeettggg	4260
aaggtgcttg	tggatatttt	ggcaggttat	ggagcagggg	tggcaggcgc	gctcgtggcc	4320
tttaaggtga	tgagcggcga	astaccatca	accdaddacc	toottaacct	actccctcct	4380
LLLaayguda	Lyageggega	gatgetetee	accegageacc	cggccaacoo	~~~	4440
atcctctccc	ctggcgccct	agtcgtcggg	gregrgrgeg	cagcgatact	gegreggeat	
ataaacccaa	gggaggggc	tatacaataa	atgaaccggc	tgatagcgtt	cgcttcgcgg	4500
autaannan	tctccccac	acactatata	cctaaaaaca	acoctocacc	acototcact	4560
39caaccacy			23-3-2-3		ataretasea	4620
cagatcctct	ctagtcttac	carcactcag	ccyctgaaga	gycttcacca	ycygaccaac	
gaggactqct	aanaaaanta	ctccggctcg	tggctaagag	atgtttggga	ttggatatgc	4680
	ccacyccacy				-++~~~~~~	4740
acqatatta	ctgatttcaa	gacctgggtg	cagtccaage	fccfaccaca	attuccuuud	4/40
acggtgttga	ctgatttcaa	gacctggctc	cagtccaagc	teetgeegeg	coocatoato	
gtccccttct	ctgatttcaa tctcatgtca	gacctggctc acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	4800
gtccccttct	ctgatttcaa tctcatgtca	gacctggctc acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	4800 4860
gtccccttct caaaccacct	ctgatttcaa tctcatgtca gcccatgtgg	gacctggctc acgtgggtac agcacagatc	aagggagtct accggacatg	ggcggggcga tgaaaaacgg	cggcatcatg ttccatgagg	4800
gtccccttct caaaccacct atcgtggggc	ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg	gacctggctc acgtgggtac agcacagatc tagtaacacg	aagggagtct accggacatg tggcatggaa	ggcggggcga tgaaaaacgg cattccccat	cggcatcatg ttccatgagg taacgcgtac	4800 4860 4920
gtccccttct caaaccacct atcgtggggc accacgggcc	ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc	gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg	aagggagtct accggacatg tggcatggaa ccaaattatt	ggcggggcga tgaaaaacgg cattccccat ctagggcgct	cggcatcatg ttccatgagg taacgcgtac gtggcgggtg	4800 4860 4920 4980
gtccccttct caaaccacct atcgtgggc accacgggcc gctgctgagg	ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc agtacgtgga	gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg ggttacgcg	aagggagtet accggacatg tggcatggaa ccaaattatt gtgggggatt	ggcggggcga tgaaaaacgg cattccccat ctagggcgct tccactacgt	cggcatcatg ttccatgagg taacgcgtac gtggcgggtg gacgggcatg	4800 4860 4920 4980 5040
gtccccttct caaaccacct atcgtgggc accacgggcc gctgctgagg	ctgatttcaa tctcatgtca gcccatgtgg ctaggacctg cctgcacgcc	gacctggctc acgtgggtac agcacagatc tagtaacacg ctccccggcg ggttacgcg	aagggagtet accggacatg tggcatggaa ccaaattatt gtgggggatt	ggcggggcga tgaaaaacgg cattccccat ctagggcgct tccactacgt	cggcatcatg ttccatgagg taacgcgtac gtggcgggtg gacgggcatg	4800 4860 4920 4980

gatggggtgc ggttgcacag gtacgctcca gcgtgcaaac ccctcctacg ggaggaggtc 5160 acattcctgg tcgggctcaa tcaatacctg gttgggtcac agctcccatg cgagcccgaa 5220 ccggacgtag cagtgctcac ttccatgctc accgacccct cccacattac ggcggagacg 5280 gctaagcgta ggctggccag gggatetece eceteettgg ccageteate agetatecag 5340 ctgtctgcgc cttccttgaa ggcaacatgc actacccgtc atgactcccc ggacgctgac 5400 ctcatcgagg ccaacctcct gtggcggcag gagatgggcg ggaacatcac ccgcgtggag 5460 tcagaaaata aggtagtaat tttggagtct ttcgagccgc tccaagcgga ggaggatgag 5520 agggaagtat ccgttccggc ggagatcctg cggaggtcca ggaaattccc tcgagcgatg 5580 cccatatggg cacgccgga ttacaaccct ccactgttag agtcctggag ggacccggac 5640 tacgtccctc cagtggtaca cgggtgtcca ttgccgcctg ccaaggcccc tccggtacca 5700 5760 cctccacqqa qgaaqaggac ggttgtcctg ccagaatcta ccgtgtcttc cgccctggcg 5820 5880 acggcctctc ctgactagcc ctccaacgac ggcgacgcgg gatccgacgt tgagtcgtac teccecatge eccecettga gggggageeg ggggateeeg attecagega egggtettgg 5940 tctaccgtaa gcgaggggt tagtgaggac gtcgtctgct gctcgatgtc ctacacatgg 6000 acaggegece tgateacgee atgegetgeg gaggaaacea agetgeecat caatgeactg 6060 agcaactett tgeteegtea ceacaacttg gtetatgeta caacateteg cagegeaage 6120 ctgcggcaga agaaggtcac ctttgacaga ctgcaggtcc tggacgacca ctaccgggac 6180 gtgctcaagg agatgaaggc gaaggcgtcc acagttaagg ctaaacttct atccgtggag 6240 6300 gaageetgta agetgaegee eccaeatteg gecagateta aatttggeta tggggeaaag gacgtccgga acctatccag caaggccgtt aaccacatcc gctccgtgtg gaaggacttg 6360 ctggaagaca ctgagacacc aattgacacc accatcatgg caaaaaatga ggttttctgc 6420 gtccaaccag agaaggggg ccgcaagcca gctcgcctta tcgtattccc agatttgggg 6480 gttcgtgtgt gcgagaaaat ggccctttac gatgtggtct ccaccctccc tcaggccgtg 6540 6600 atgggctctt catacggatt ccaatactct cctggacagc gggtcgagtt cctggtgaat 6660 gcctggaaag cgaagaaatg ccctatgggc ttcgcatatg acacccgctg ttttgactca 6720 acqqtcactq aqaatqacat ccqtqttgag gagtcaatct accaatqttg tgacttggcc 6780 cccgaagcca gacaggccat aaggtcgctc acagagcggc tttacatcgg gggccccctg 6840 actaattcta aagggcagaa ctgcggctat cgccggtgcc gcgcgagcgg tgtactgacg accagctgcg gtaataccct cacatgttac ttgaaggccg ctgcggcctg tcgagctgcg 6900 aagctccagg actgcacgat gctcgtatgc ggagacgacc ttgtcgttat ctgtgaaagc 6960 gcggggaccc aagaggacga ggcgagccta cgggccttca cggaggctat gactagatac 7020 7080 tctgccccc ctggggaccc gcccaaacca gaatacgact tggagttgat aacatcatgc 7140 tcctccaatg tgtcagtcgc gcacgatgca tctggcaaaa gggtgtacta tctcacccgt gaccccacca cccccttgc gcgggctgcg tgggagacag ctagacacac tccagtcaat 7200 7260 tcctggctag gcaacatcat catgtatgcg cccaccttgt gggcaaggat gatcctgatg 7320 actcatttct tctccatcct tctagctcag gaacaacttg aaaaagccct agattgtcag 7380 atctacgggg cctgttactc cattgagcca cttgacctac ctcagatcat tcaacgactc catggcctta gcgcattttc actccatagt tactctccag gtgagatcaa tagggtggct 7440 7500 tcatgcctca ggaaacttgg ggtaccgccc ttgcgagtct ggagacatcg ggccagaagt gtccgcgcta ggctactgtc ccaggggggg agggctgcca cttgtggcaa gtacctcttc 7560 aactgggcag taaggaccaa gctcaaactc actccaatcc cggctgcgtc ccagttggat 7620 ttatccagct ggttcgttgc tggttacagc gggggagaca tatatcacag cctgtctcgt 7680 gcccgacccc gctggttcat gtggtgccta ctcctacttt ctgtaggggt aggcatctat 7740 7800 ctactcccca accgatgaac ggggacctaa acactccagg ccaataggcc atcctgtttt 7860 7920 tttttcctct ttttttcctt ttctttcctt tggtggctcc atcttagccc tagtcacggc tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc 7980 agatcaagta ctcctgcagg cgcgccacta gtgggaatac gcggggtatg ccgcgtttta 8040 gcatattgac gacccaattc tcatgtttga cagcttatca tcgataagct ttaatgcggt 8100 agtttatcac agttaaattg ctaacgcagt caggcaccgt gtatgaaatc taacaatgcg 8160 8220 ctcatcgtca tecteggeac cgtcaccctg gatgetgtag gcataggett ggttatgecg gtactgccgg gcctcttgcg ggatatcgtc cattccgaca gcatcgccag tcactatggc 8280 8340 gtgctgctag cgctatatgc gttgatgcaa tttctatgcg cacccgttct cggagcactg 8400 tecgaeeget ttggeegeeg eeeagteetg etegettege taettggage eaetategae tacgcgatca tggcgaccac acccgtcctg tggatcctct acgccggacg catcgtggcc 8460 ggcatcaccg gcgccacagg tgcggttgct ggcgcctata tcgccgacat caccgatggg 8520 8580 gaagatcggg ctcgccactt cgggctcatg agcgcttgtt tcggcgtggg tatggtggca 8640 ggccccgtgg ccgggggact gttgggcgcc atctccttgc atgcaccatt ccttgcggcg 8700 gcggtgctca acggcctcaa cctactactg ggctgcttcc taatgcagga gtcgcataag ggagagegte gacegatgee ettgagagee tteaacecag teageteett ceggtgggeg 8760 cggggcatga ctatcgtcgc cgcacttatg actgtcttct ttatcatgca actcgtagga 8820 8880 caggtgccgg cagcgctctg ggtcattttc ggcgaggacc gctttcgctg gagcgcgacg 8940 atgateggee tgtegettge ggtattegga atettgeaeg ceetegetea ageettegte actggtcccg ccaccaaacg tttcggcgag aagcaggcca ttatcgccgg catggcggcc 9000 9060 gacgcgctgg gctacgtctt gctggcgttc gcgacgcgag gctggatggc cttccccatt

atgattcttc	tcgcttccgg	cggcatcggg	atgcccgcgt	tgcaggccat	gctgtccagg	9120
caggtagatg	acgaccatca	gggacagctt	caaggatcgc	tegeggetet	taccagccta	9180
acttcgatca	ctggaccgct	gatcgtcacg	gcgatttatg	ccgcctcggc	gagcacatgg	9240
aacgggttgg	catggattgt	aggcgccgcc	ctataccttg	tatgaataaa	cgcgttgcgt	9300
cgcggtgcat	ggagccgggc	cacctcgacc	tgaatggaag	ccggcggcac	ctcgctaacg	9360
gattcaccac	tccaagaatt	ggagccaatc	aattcttgcg	gagaactgtg	aatgcgcaaa	9420
ccaacccttq	gcagaacata	tccatcgcgt	ccgccatctc	cagcagccgc	acgcggcgca	9480
tctcgggcag	cgttgggtcc	tggccacggg	tgcgcatgat	cgtgctcctg	tcgttgagga	9540
cccqqctaqq	ctggcggggt	tgccttactg	gttagcagaa	tgaatcaccg	atacgcgagc	9600
gaacgtgaag	cgactgctgc	tgcaaaacgt	ctgcgacctg	agcaacaaca	tgaatggtct	9660
tcqqtttccq	tgtttcgtaa	agtctggaaa	cgcggaagtc	agcgccctgc	accattatgt	9720
tccqqatctq	catcgcagga	tgctgctggc	taccctgtgg	aacacctaca	tctgtattaa	9780
cgaagcgctg	gcattgaccc	tgagtgattt	ttctctggtc	ccgccgcatc	cataccgcca	9840
gttgtttacc	ctcacaacgt	tccagtaacc	gggcatgttc	atcatcagta	acccgtatcg	9900
tgagcatcct	ctctcgtttc	atcggtatca	ttacccccat	gaacagaaat	tececettae	9960
acqqaqqcat	caagtgacca	aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
aaqccaqaca	ttaacgcttc	tggagaaact	caacgagctg	gacgcggatg	aacaggcaga	10080
catctqtqaa	tcgcttcacg	accacgctga	tgagctttac	cgcagctgcc	tegegegttt	10140
cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	cagcttgtct	10200
gtaagcggat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	10260
tcqqqqcqca	gccatgaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	10320
gcggcatcag	agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tgcgtaagga	qaaaataccg	catcaggcgc	tcttccgctt	cctcgctcac	tgactcgctg	10440
cacteagteg	ttcggctgcg	gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat	caqqqqataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	10740
ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	10800
aggtatctca	gttcggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	10860
gttcagcccg	accgctgcgc	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	10920
cacqacttat	cgccactggc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	10980
ggcggtgcta	cagagttctt	gaagtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttqqtatct	gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	11100
tccggcaaac	aaaccaccqc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
cqcaqaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	11220
tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
tagatccttt	tctagataat	acgactcact	ata			11313

```
<210> 7
<211> 11313
<212> DNA
<213> Artificial Sequence
<220>
<223> Plasmid
```

### <400> 7 /

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 120 tetteacgca gaaagcgtet agccatggcg ttagtatgag tgtcgtgcag cetecaggae 180 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 240 gacgaccggg tectttettg gateaacccg eteaatgeet ggagatttgg gegtgeecce gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360 420 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 480 cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540 acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600 660 cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 720 tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 780 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg agcgagcacg tactcggatg gaagccggtc 840 900 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 960 ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct

gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1020 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080 ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140 agegeatege ettetatege ettettgaeg agttettetg agtttaaaca gaecacaacg 1200 gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260 cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1320 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct 1380 aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440 gttcctctgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1500 aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560 1620 gcaaaggegg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa 1680 tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggt accccattgt 1740 atgggatetg atetggggee teggtgeaca tgetttaeat gtgtttagte gaggttaaaa aacgtctagg ccccccgaac cacggggacg tggttttcct ttgaaaaaca cgataatacc 1800 atggcgccta ttacggccta ctcccaacag acgcgaggcc tacttggctg catcatcact 1860 agceteacag geegggacag gaaccaggte gagggggagg tecaagtggt etceacegea 1920 acacaatett teetggegae etgegteaat ggegtgtgtt ggaetgteta teatggtgee 1980 ggctcaaaga cccttgccgg cccaaagggc ccaatcaccc aaatgtacac caatgtggac 2040 2100 caggaceteg teggetggea agegeeece ggggegegtt cettgacace atgeacetge 2160 ggcagcgcgg acctttactt ggtcacgagg catgccgatg tcattccggt gcgccggcgg ggcgacagca gggggagcct actctccccc aggcccgttt cctacttgaa gggctcttcg 2220 ggcggtccac tgctctgccc ctcggggcac gctgtgggca tctttcgggc tgccgtgtgc 2280 acccgagggg ttgcgaaggc ggtggacttt gtacccgtcg agtctatggg aaccactatg 2340 2400 eggteeeegg tetteaegga caactegtee ceteeggeeg tacegeagae atteeaggtg 2460 qcccatctac acqcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 2520 gcccaagggt ataaggtgct tgtcctgaac ccgtccgtcg ccgccaccct aggtttcggg 2580 gcqtatatgt ctaaggcaca tggtatcgac cctaacatca gaaccggggt aaggaccatc 2640 accaegggtg eccecateae gtactecace tatggeaagt ttettgeega eggtggttge 2700 tctgggggcg cctatgacat cataatatgt gatgagtgcc actcaactga ctcgaccact atcctgggca tcggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtcgtg 2760 ctogccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggaggtg 2820 2880 gctctgtcca gcactggaga aatccccttt tatggcaaag ccatccccat cgagaccatc aaggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccgcg 2940 3000 aagetgteeg geeteggaet caatgetgta geatattace ggggeettga tgtateegte ataccaacta geggagaegt cattgtegta geaacggaeg etetaatgae gggetttace 3060 ggcgatttcg actcagtgat cgactgcaat acatgtgtca cccagacagt cgacttcagc 3120 ctggacccga ccttcaccat tgagacgacg accgtgccac aagacgcggt gtcacgctcg 3180 3240 cagcggcgag gcaggactgg taggggcagg atgggcattt acaggtttgt gactccagga gaacggccct cgggcatgtt cgattcctcg gttctgtgcg agtgctatga cgcgggctgt 3300 3360 gettggtacg agetcaegee egeegagaee teagttaggt tgegggetta cetaaacaca ccagggttgc ccgtctgcca ggaccatctg gagttctggg agagcgtctt tacaggcctc 3420 acccacatag acgcccattt cttgtcccag actaagcagg caggagacaa cttcccctac 3480 ctggtagcat accaggctac ggtgtgcgcc agggctcagg ctccacctcc atcgtgggac 3540 3600 caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gcccctgctg tataggctgg gagccgttca aaacgaggtt actaccacac accccataac caaatacatc 3660 3720 atggcatgca tgtcagctga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 3780 gtcctagcag ctctggccgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg atcatcttgt ccggaaagcc ggccatcatt cccgacaggg aagtctttta ccgggagttc 3840 gatgagatgg aagagtgcgc ctcacacctc ccttacatcg aacggggaat gcagctcgcc 3900 3960 gaacatttca aacagaaggc aatcgggttg ctgcaaacag ccaccaagca agcggaggct 4020 gctgctcccg cggtggaatc caagtggcgg accctcgaag ccttctgggc gaagcatatg 4080 tqqaatttca tcaqcqqqat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4140 gcgatagcat cactgatggc attcacagcc tctatcacca gcccgctcac cacccaacat 4200 accetectgt ttaacatect ggggggatgg gtggccgccc aacttgctcc tcccagcgct 4260 gettetgett tegtaggege eggeateget ggageggetg ttggeageat aggeettggg aaggtgcttg tggatatttt ggcaggttat ggagcagggg tggcaggcgc gctcgtggcc 4320 tttaaggtca tgagcggcga gatgccctcc accgaggacc tggttaacct actccctgct 4380 atcetetece etggegeet agtegteggg gtegtgtgeg eagegataet gegteggeae 4440 4500 gtgggcccag gggaggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgcgg 4560 ggtaaccacg tctcccccac gcactatgtg cctgagagcg acgctgcagc acgtgtcact cagatectet etagtettae cateacteag etgetgaaga ggetteacea gtggateaac 4620 gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatatgc 4680 4740 acggtgttga ctgatttcaa gacctggctc cagtccaagc tcctgccgcg attgccggga 4800 gtccccttct tctcatgtca acgtgggtac aagggagtct ggcggggcga cggcatcatg caaaccacct gcccatgtgg agcacagatc accggacatg tgaaaaacgg ttccatgagg 4860 4920 ategtggggc ctaggacctg tagtaacacg tggcatggaa cattccccat taacgcgtac

accacaaaacc	cctgcacgcc	chacadada	ccaaattatt	ctagggcgct	ataacaaata	4980
	agtacgtgga	cattoaggag	atagagaatt	tagactacat	dacadadata	5040
gergergagg	agcacgugga	ggctacgcgg	gragagarr	ccaccacge	gacgggcacg	
accactgaca	acgtaaagtg	cccgtgtcag	grrccggccc	ccgaattett	cacagaagtg	5100
gatggggtgc	ggttgcacag	gtacgctcca	gcgtgcaaac	ccctcctacg	ggaggaggtc	5160
acattcctgg	tcgggctcaa	tcaatacctg	gttgggtcac	agctcccatg	cgagcccgaa	5220
acacacataa	cagtgctcac	ttccatoctc	accdacccct	cccacattac	aacaaaaaca	5280
ccggacgcag	Cagigordac	tetaca	accgacccc	agagatasta	33033223	5340
gctaagcgta	ggctggccag	gggatettee	ecereering	ceagercare	agctagctag	
ctgtctgcgc	cttccttgaa	ggcaacatgc	actacccgtc	atgactcccc	ggacgctgac	5400
ctcatcgagg	ccaacctcct	gtggcggcag	gagatgggcg	ggaacatcac	ccgcgtggag	5460
traraaaata	aggtagtaat	tttggactct	ttcgagccgc	tccaagcgga	ggaggatgag	5520
	aggrageaac	cccggacccc	accacataca	aceasttaca	tagarata	5580
agggaagtat	ccgttccggc	ggagaccccy	eggaggteea	ggaaarreec	ccgagcgacg	
cccatatggg	cacgcccgga	ttacaaccct	ccactgttag	agreerggag	ggacccggac	5640
tacqtccctc	cagtggtaca	cgggtgtcca	ttgccgcctg	ccaaggcccc	tccggtacca	5700
cctccacgga	ggaagaggac	aattateeta	ccagaatcta	ccqtqtcttc	caccctaaca	5760
asaataaas	caaagacctt	caacsactcc	gaatcgtcgg	ccat.caacaa	caacacaaca	5820
gagetegeta	caaagacccc	character	2446026023	astagasast	tasataatsa	5880
acggcctctc	ctgactagcc	CLCCaacgac	ggcgacgcgg	yaccegacge	LyayLuguau	
tcccccatgc	cccccttga	gggggagccg	ggggatcccg	attccagcga	cgggtcttgg	5940
tctaccqtaa	gcgagggggt	tagtgaggac	gtcgtctgct	gctcgatgtc	ctacacatgg	6000
acaddddddd	tgatcacgcc	atacactaca	gaggaaacca	agctgcccat	caatgcactg	6060
202330300	tgctccgtca	acadaactta	atatatacta	caacatctco	cadcacaaac	6120
agcaactctt	Lycicogica	ccacaaccig	guccauguca		attagggaage	6180
ctgcggcaga	agaaggtcac	ctttgacaga	ctgcaggtcc	tggacgacca	ccaccyyyac	
gtgctcaagg	agatgaaggc	gaaggcgtcc	acagttaagg	ctaaacttct	atccgtggag	6240
gaageetgta	agctgacgcc	cccacattcq	gccagatcta	aatttggcta	tggggcaaag	6300
2 datadas	acctatccag	caannoontt	aaccacatco	acticatata	gaaggacttg	6360
gatgtttgga	accuaticag	caaggeegee	accuoatco	9000090909	gattttataa	6420
ctggaagaca	ctgagacacc	aattgacacc	accardatgg	Caaaaaacya	ggcccccgc	4*
gtccaaccag	agaagggggg	ccgcaagcca	gctcgcctta	tcgtattccc	agatttgggg	6480
attcatatat	gcgagaaaat	ggccctttac	gatgtggtct	ccaccctccc	tcaggccgtg	6540
atogoctctt	catacggatt	ccaatactct	cctggacagc	agatcaaatt	cctqqtqaat	6600
acgggcccc	cgaagaaatg	acatataaa	ttcccatatc	acacccacta	ttttgactca	6660
gccrygaaay	cgaagaaatg	ccccacgggc		acaccegeeg	tanattaaaa	6720
acggtcactg	agaatgacat	cegtgttgag	gagteaatet	accaarging	Lyactiggee	
cccgaagcca	gacaggccat	aaggtcgctc	acagagcggc	tttacatcgg	gggccccctg	6780
actaattcta	aagggcagaa	ctgcggctat	caccaatacc	gcgcgagcgg	tgtactgacg	6840
accadetded	gtaataccct	cacatottac	ttgaaggcgg	ctacaaccta	tcgagctgcg	6900
accagoogog	gedaedcoot	actactatac	adadadada	ttatcattat	ctatassac	6960
aagctccagg	actgcacgat	geregrange	ggagacgacc	cogcogciac	cegegaaage	
gcggggaccc	aagaggacga	ggcgagccta	cgggccttca	cggaggctat	gactagatac	7020
tetgecece	ctggggaccc	gcccaaacca	gaatacgact	tggagttgat	aacatcatgc	7080
tcctccaatq	tgtcagtcgc	gcacgatgca	tctqqcaaaa	gggtgtacta	tctcacccgt	7140
222222222	cccccttgc	acadactaca	taggagagag	chagadadad	tccagtcaat	7200
gaccccacca		9099900909	cagagagacag	aaaasaast	astactasta	7260
teetggetag	gcaacatcat	catgtatgcg	Cocaccity	gggcaaggac	gaccccgacg	7320
actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaaagccct	agattgtcag	
atctacqqqq	cctgttactc	cattgagcca	cttgacctac	ctcagatcat	tcaacgactc	7380
catggcctta	gcgcattttc	actccatagt	tactctccag	qtqaqatcaa	tagggtggct	7440
tastaatas	ggaaacttgg	autaceace	ttacaaatct	ggagagatgg	aaccaaaaat.	7500
Leatycetta	ggaaacttgg	ggcaccgccc	cogogageee	ggagacacog	atacatatta	7560
gtccgcgcta	ggctactgtc	ccaggggggg	agggergeea	Crrdradcaa	gracectore	
aactgggcag	taaggaccaa	gctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	7620
ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	7680
decedacee	gctggttcat	ataataccta	ctcctacttt	ctgtaggggt	aggcatctat	7740
gtagtagaga	accgatgaac	gagagataa	acactccagg	ccaataggc	atcctqtttt	7800
CLacicocca	accyacyaac	9999acccaa	*****	++++++++	ttataattt	7860
tttccctttt	ttttttttt	TTTTTTTTT			to a state of the	
tttttcctct	ttttttcctt	ttctttcctt	tggtggctcc	atcttagccc	tagtcacggc	7920
tagctgtgaa	aggtccgtga	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	7980
agatcaagta	ctcctgcagg	cococcacta	gtgggaatac	gcggggtatg	ccgcgtttta	8040
gatouagea	gacccaattc	tastatttas	caccttatca	troataaoct	ttaatgcggt	8100
geatattgae	gacccaaccc	ccatgettga	cagoccacca	-t-t	tananataaa	8160
agtttatcac	agttaaattg	ctaacgcagt	caggeacege	grargaaarc	caacaacycy	
ctcatcgtca	tcctcggcac	cgtcaccctg	gatgctgtag	gcataggctt	ggttatgccg	8220
gtactqccqq	gcctcttgcg	ggatatcgtc	cattccgaca	gcatcgccag	tcactatggc	8280
atactacted	cgctatatgc	attaatacaa	tttctataca	cacccattct	cggaqcactq	8340
505050009	ttggccgccg		ctcccttccc	tacttocacc	cactaticgae	8400
Lougaucget	Lugguuguug		transt = =t ='	2222232	datactac-	8460
tacgcgatca	tggcgaccac	accegteetg	iggarceret	acgccggacg	carcyryyce	
ggcatcaccg	gcgccacagg	tgcggttgct	ggcgcctata	tcgccgacat	caccgatggg	8520
gaagatcggg	ctcgccactt	cgggctcatg	agcgcttgtt	teggegtggg	tatggtggca	8580
~~~~~~~~~~~	_	~======================================	atctccttcc	atocaccatt	ccttqcqqca	8640
	- acadadaaaa	OFFIGURES:				
ggccccgcgg	ccgggggact	grtgggegee	aaataattaa	taatooaooa	atcacatasa	ጸማሰበ
gcggtgctca	acggcctcaa	cctactactg	ggctgcttcc	taatgcagga	gtcgcataag	8700
gcggtgctca ggagagcgtc	acggcctcaa gaccgatgcc	cctactactg cttgagagcc	ggctgcttcc ttcaacccag	taatgcagga tcagctcctt	gtcgcataag ccggtgggcg	8760
gcggtgctca ggagagcgtc cggggcatga	acggcctcaa gaccgatgcc ctatcgtcgc	cctactactg cttgagagcc cgcacttatg	ggctgcttcc ttcaacccag actgtcttct	taatgcagga tcagctcctt ttatcatgca	gtcgcataag ccggtgggcg actcgtagga	8760 8820
gcggtgctca ggagagcgtc cggggcatga	acggcctcaa gaccgatgcc	cctactactg cttgagagcc cgcacttatg	ggctgcttcc ttcaacccag actgtcttct	taatgcagga tcagctcctt ttatcatgca	gtcgcataag ccggtgggcg actcgtagga	8760

atgatcggcc	tatcacttac	ggtattcgga	atcttgcacg	ccctcgctca	agccttcgtc	8940
actootecco	ccaccaaacq	tttcggcgag	aagcaggcca	ttatcgccgg	catggcggcc	9000
aacacactaa	actacatett	actagagtta	gcgacgcgag	gctggatggc	cttccccatt	9060
atgattette	tcacttccaa	cggcatcggg	atgcccgcgt	tgcaggccat	gctgtccagg	9120
caggtagatg	acqaccatca	gggacagctt	caaggatcgc	tegeggetet	taccagccta	9180
acttcgatca	ctggaccgct	gatcgtcacg	gcgatttatg	ccgcctcggc	gagcacatgg	9240
aacqqqttqq	catggattgt	aggcgccgcc	ctataccttg	tatgactaca	cgcgttgcgt	9300
cacaatacat	ggagccgggc	cacctcgacc	tgaatggaag	ccggcggcac	ctcgctaacg	9360
gattcaccac	tccaagaatt	ggagccaatc	aattcttgcg	gagaactgtg	aatgcgcaaa	9420
ccaacccttq	gcagaacata	tccatcgcgt	ccgccatctc	cagcagccgc	acgcggcgca	9480
tctcgggcag	cgttgggtcc	tggccacggg	tgcgcatgat	cgtgctcctg	tcgttgagga	9540
cccqqctaqq	ctqqcggggt	tgccttactg	gttagcagaa	tgaatcaccg	atacgcgagc	9600
gaacgtgaag	cgactgctgc	tgcaaaacgt	ctgcgacctg	agcaacaaca	tgaatggtct	9660
tcggtttccg	tqtttcqtaa	agtctggaaa	cgcggaagtc	agcgccctgc	accattatgt	9720
tccqqatctq	catcgcagga	tgctgctggc	taccctgtgg	aacacctaca	tctgtattaa	9780
cgaagcgctg	gcattgaccc	tgagtgattt	ttctctggtc	ccgccgcatc	cataccgcca	9840
gttgtttacc	ctcacaacgt	tccagtaacc	gggcatgttc	atcatcagta	acccgtatcg	9900
tgagcatcct	ctctcgtttc	atcggtatca	ttacccccat	gaacagaaat	tcccccttac	9960
acggaggcat	caagtgacca	aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
aaqccaqaca	ttaacqcttc	tggagaaact	caacgagctg	gacgcggatg	aacaggcaga	10080
catctgtgaa	tcgcttcacg	accacgctga	tgagctttac	cgcagctgcc	tcgcgcgttt	10140
cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	cagcttgtct	10200
gtaagcggat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	10260
teggggegea	gccatgaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	10320
gcggcatcag	agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tqcqtaaqqa	gaaaataccg	catcaggcgc	tetteegett	cctcgctcac	tgactcgctg	10440
cgctcggtcg	ttcggctgcg	gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat	caggggataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	10740
ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	10800
aggtatctca	gttcggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	10860
gttcagcccg	accgctgcgc	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	10920
cacqacttat	cgccactggc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	10980
ggcggtgcta	cagagttctt	gaagtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttqqtatct	gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	tagçtcttga	11100
tccqqcaaac	aaaccaccgc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
cqcaqaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	11220
tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
	tctagataat					11313

```
<210> 8
<211> 11313
<212> DNA
<213> Artificial Sequence
<220>
<223> Plasmid
```

### <400> 8

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 120 tetteaegea gaaagegtet ageeatggeg ttagtatgag tgtegtgeag eeteeaggae 180 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 240 gacgaccggg teettettg gatcaacccg etcaatgcet ggagatttgg gegtgeeece 300 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 360 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 420 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480 ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540 600 acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660 tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720 780 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc

tegtogataa ggegegeatg cegaeggeg agatectog eggeatgeec gegestgeec 960 cetagegea atteatggtg gaaatggee getteteeg tegtetteeg atteategge deggetgeec eeggeatgeec 1080 teggegegeatgeec tettetaege ettetteeg tegtetteegg atteategg ageogatege cetteteeg tegtetteegg ageogatege eggetgeet atteogate eggetgeet eeggeatgeec tetteatge ageogatege ettetteeg agettetteeg agetteategg ageogatege tetteatge deggeatge atteogate atteogate ageogatege tetteatge ageogate atteogate atteogate ageogate ag							
tetgocaga gagateteg gagateteg cytagogag gagateteg cytagogag gagateged yespoggagag gagateged yespoggagag gagateged yespoggagag gagateged yespoggagag gagateged yespoggagag gagateged yespoggagagagagagagagagagagagagagagagagagag	cattcqacca	ccaagcgaaa	catcgcatcg	agcgagcacg	tacteggatg	gaagccggtc	840
ccaggactaaa gacagcatat cadagagga gagatcteg cyttycogaa tateataggga gaacagtat cadagagga atsegatgga gaacagcat cadagagactaa gattatatag tateatagga cyttotectog tyttectetaa gacagataaa gacagataaa gattatataa gattataaaa gacaaaaaa cagtatagaa gattatataa gacaaaaaaa cagtagataaaa gacacaatag tateatagaa aacacacaaaa aagtatataaa gattatataa gacaaaaaaa aagtatataaaaaaa aagtatattaaa gacaaaaaaa aagtatataaaaaaa aagtatataaaaaaaa	ttataantan	aastastata	uscussusuc	atcagggggt	cacaccaacc	gaactgttcg	900
gottetic gagacaga atgagetgas ogategate ogtteteting attactogac bytogagaga atgagetgas ogategates of otteteting agottetetic agottetetic agotteteting attropped telescope of teles	Ligitigatea	ggargarcing	gacgaagage	22025550	cgcgccagcc	3	0.00
gottetic gagacaga atgagetgas ogategate ogtteteting attactogac bytogagaga atgagetgas ogategates of otteteting agottetetic agottetetic agotteteting attropped telescope of teles	ccaqqctcaa	ggcgcgcatg	cccgacggcg	aggatetegt	cgtgacccat	ggcgatgcct	960
tgggtqatagg ggaacgetat caggacatag cgttsggtac cettgatatt gctgaagaag 1988 ttaggaggaa atgggtgata cggtttactet ggttatatag atacgecget cogatatag gtttectet agagagaaga cuttettagg attettatag attegacgat ggtttactet agagagaaga caggacqut ttggaatagg gacacaaaga ggagaactag ggttattatt gaaattggaagaga ggaaagaagaagaagagagag	acttaccass	tatcatooto	gaaaatggcc	acttttctaa	attcatcgac	tataaccaac	1020
tegogacoga atgogacoga estetatedo estetatedo general estato de estato est	goregeegaa		January 3300			actassasa	
gegotatogo ottotatogo descriptiones altoegococo ottocotoco occocotaa ogitacidego 1200 gegotatogo tidgadatago ogotottoto godototto occocotaa ogitacidego 1320 gegotogoto tidgadatago ogototototogo aagototot tidgadatago ogototototogo aagotocococaa ottogodogo godotototogo godotogodotototogo aagotocococaa ottogodogo ogotococococococococococococococococococ	tgggtgtggc	ggaccgctat	caggacatag	egriggerae	Cegigalati	gergaagage	
gegotatogo ottotatogo descriptiones altoegococo ottocotoco occocotaa ogitacidego 1200 gegotatogo tidgadatago ogotottoto godototto occocotaa ogitacidego 1320 gegotogoto tidgadatago ogototototogo aagototot tidgadatago ogototototogo aagotocococaa ottogodogo godotototogo godotogodotototogo aagotocococaa ottogodogo ogotococococococococococococococococococ	ttaacaacaa	atoggctgac	cacttcctca	tgctttacgg	tatcgccgct	cccgattcgc	1140
getteccete agoggaatea atteogece tetecetece cocceccaa ogtacotgga togaaacoget togaaataag coggstaged tettettaa tytattite caccatatig 1380 agoggstatt coccutetored caaaggagate caaagatett gaagtotg gagoggaate caagactett gaagtotg agoggaagaga agoggaagagagaacaacaaccaa gagogaacacaacaacaacaacaacaacaacacaac	55-55-5-	-t-c+c+ccc	attattaaaa	agttettete	actttaaaca	uaccacaacu	1200
gagagocgat tggaataagg coggtgtggg tittgtctata tgttatittd caccatattg cagatgtgag gacogaaa caggcctg tettettiga gagaatett gaagagatgg aggettett gaagagaatg caaggtctgt tgaatgtcg gagaattct gagagatcg caccatcac gtgccaccattg gagcattgga tggaattctg gaagcaccacca gtgccaccacca gtgcaccattg gaccaccttg aggaagagag caccaccacca gtgccaccat gtgaatgga cacggtgta aggacaccatt gaggaagaga caccaccat gtgcttataca tggtttagtga aggacacgg cacggagacg caccaccacca gtgccacacacca tggtttagaa tagttgtgga acccattgt 1680 aacgctatag caccaccacca gtgccacacacca tggtttacat gtgtttagat acccattgt 1680 aacgctcaag caccaccac cacacggagagag tggtttect ttgaaaaaca cgataataca cacacacaca cacacaca caccacacac cacacaca geggagaggg ttgtttect ttgaaaaaca cgataatacc accacacaca caccacacacacac cacacaca	agegearege	CLLCLatege	Culculgacy	ageceeeeg	ageceaaaca	gaocacaacg	
gagagocgat tggaataagg coggtgtggg tittgtctata tgttatittd caccatattg cagatgtgag gacogaaa caggcctg tettettiga gagaatett gaagagatgg aggettett gaagagaatg caaggtctgt tgaatgtcg gagaattct gagagatcg caccatcac gtgccaccattg gagcattgga tggaattctg gaagcaccacca gtgccaccacca gtgcaccattg gaccaccttg aggaagagag caccaccacca gtgccaccat gtgaatgga cacggtgta aggacaccatt gaggaagaga caccaccat gtgcttataca tggtttagtga aggacacgg cacggagacg caccaccacca gtgccacacacca tggtttagaa tagttgtgga acccattgt 1680 aacgctatag caccaccacca gtgccacacacca tggtttacat gtgtttagat acccattgt 1680 aacgctcaag caccaccac cacacggagagag tggtttect ttgaaaaaca cgataataca cacacacaca cacacaca caccacacac cacacaca geggagaggg ttgtttect ttgaaaaaca cgataatacc accacacaca caccacacacacac cacacaca	gtttccctct	agcgggatca	attccgcccc	tetecetece	cccccctaa	cgttactggc	T200
cegtettittig geaatgtaga ggeceggaaa ectggecetg tettettigae gagsatiect gagsgastiet coccetetoge eaaggastieg tagaatgtedg taaatgtedg aagsgaageag 1440 gaccectetg acacgeaage decetetoge gecaaaage egtgetata gaagacaaga gtgeetetotge gecaaaage egtgetata gaagatacaaa 1520 geaaaggegg eaacaaceca gtgecaegt gtgattigat gaagatgtag aagatacaact 1560 geaaaggegg eacaaggegg et eaggggatet eaacagggg etgaagatg eagtgttaga aagagteaaa 1620 aagggatett eaacaagggg etgaagatg eagtgttaga aagagteaaa 1620 aagggatett eacaagggg etgaaggagg etgtttagt gaacaagatg eacaagagt eacaagagga etgaaaaacatt teetggegae etgegtaaa gageggagg gagegggg eacaagggg eacaagggg eacaagggg eacaagggggggggg	cassaccact	tagaataaga	regatataca	tttgtctata	tottattttc	caccatatto	1320
aggsgtettt ocotetoge caaaggaatg caaggtotgt tgaatgtogt gaaggaagga gactotcttg agacacttttg aagcacactttg aaggaactg 1500 gacacacccac ctggcgacag gtgcotctg ggccaaaagg cacgtgtata agatacact 1560 geaaagggg caaaagggg caaaagggg caaaagggg caaaagggg caaaagggg caaaagggg caaaagggg caaaagggg caaaagggg cacacacca gtgccacagtt gtgattgga tattgttgga aaggttaaa 1620 accgatagtgt atcggggact cattggggac cacacttgt tatgggcac accgacaggac gacggggacag gacgggggggg	cgaageegee	cggaacaagg	0055050505		tattattaaa	anaaattaat	
actoccocca chagogacaag stycocctycy gocaaaago cactybataa gaatacacct 1560 geaaagogg caaaacoca gtyccacqtt gtsgsttyga tagttytyga aagatacacct 1560 geaaagogg caaaacoca gtyccacqtt gtsgsttyga tagttytyga aagatacacat 1560 geaaagogg caaaacoca gtyccacqtt gtsgsttyga tagttytyga aagatacacat 1560 geaaagogg cacaaacoca gtyccacquy gtsgttttact gtsgttaga aagatacacaa attyggcota actyggogaca cycgytagaca gygtttacat gygtttaga aagatacacaa aatyggacata tacagogacta tacagogacaa cocastygagaca gaggaggog tactatygyt cacaacaacactt toctygoga cacaacacact toctygoga cycacaacaagogac caatacacc aaaatytacac caaaatytacaagoctaa cocttycoga cacaaagogac cocaatcacca aaaatytacac caaatytacagagacacagagacacagagacacagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagagacacagagagagacacagagagagagacacaagagagagacacaagagagagagacacaagagagagagagagagacacaagagagagagagagagagagagagagagagagagagaga	ccgtcttttg	gcaatgtgag	ggcccggaaa	deragedera	LULLULLYAU	gageactect	
actoccocca chagogacaag stycocctycy gocaaaago cactybataa gaatacacct 1560 geaaagogg caaaacoca gtyccacqtt gtsgsttyga tagttytyga aagatacacct 1560 geaaagogg caaaacoca gtyccacqtt gtsgsttyga tagttytyga aagatacacat 1560 geaaagogg caaaacoca gtyccacqtt gtsgsttyga tagttytyga aagatacacat 1560 geaaagogg cacaaacoca gtyccacquy gtsgttttact gtsgttaga aagatacacaa attyggcota actyggogaca cycgytagaca gygtttacat gygtttaga aagatacacaa aatyggacata tacagogacta tacagogacaa cocastygagaca gaggaggog tactatygyt cacaacaacactt toctygoga cacaacacact toctygoga cycacaacaagogac caatacacc aaaatytacac caaaatytacaagoctaa cocttycoga cacaaagogac cocaatcacca aaaatytacac caaatytacagagacacagagacacagagacacagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagacacagagagagacacagagagagacacagagagagagacacaagagagagacacaagagagagagacacaagagagagagagagagacacaagagagagagagagagagagagagagagagagagagaga	aggggtcttt	cccctctcqc	caaaqqaatq	caaggtctgt	tgaatgtcgt	gaaggaagca	1440
accocceae chagogacaa glocototog ogocaaaaagc cacytytata agatacacct glocaaagogg cacaaaccca glococacyt ybgaqtsya tattytyga aaagatcaaa 1620 acagaagatcy acceayaaggt acceayaaggt acceayaaggt cacaggacta acceatty cacaggacaa cacaggactaa tataggacta cacaggaga glogaaaggag tygttaagt gagyttaaaa 1740 acagtotaag cacoggacaa gaacaagt gagygaggag tattteet theyaaaaaca cyataatacc 1800 agoctacaaa geocgaaaag cacaagga gacaagga gagygagga tattygatty gacatatacct acaagcacaacactatt tectoggaca cacaggag gagygagga gagytyty gacatatacct acaagtyga cactycoggac caggacacag cacagacagga cactycog cacaaggga cacaggac gagygagact cattygacac agocagacagaa cacttact gybacagga gagyagactacacc aaatytacac acaagtygac acctyped acttact gybacagag cacagagagacagga cactyped acttacacca agocagagagagagacaggagacaggagacacaggagagacacaggagagacacaggagagacacaggagagacacaggagagacacaggagagacacaggagagagacacaggagagagagagagagagagagagagagagagagagaga	attaatataa	22444444	224242242	acatetataa	caaccettta	caddcadcad	1500
goaaagogg cacaaccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa 1630 tgggetctoc caagoggtatt cacaagggg ctgaaggatg caccaattg 1680 aacgetctag caccoqaac cagggacg tggttttect ttgaaaaca cgataataac 1840 aacgetcag caccoqaac cacgggacg tggttttect ttgaaaaca cgataataac 1840 aacgetcaag caccoqaac cacgggacg gascacggtt ttgaaagagg ctccacacg agccggacg taccacggacg ttggggacg cacaagggg gascacgagg gaggaggg tcacaaggg cacaagggg cacaaggggg cacaaggggg cacaaggggg cacaagggg gagggggg tactataggg gagggggg cacaaggggg gagggggg tcacaagggg cacaagggg gagggggggggg	gricerergg	aayculuug	aagacaaaca	acgeeegeag	cgaccccccg	caggeagegg	
goaaagogg cacaaccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa 1630 tgggetctoc caagoggtatt cacaagggg ctgaaggatg caccaattg 1680 aacgetctag caccoqaac cagggacg tggttttect ttgaaaaca cgataataac 1840 aacgetcag caccoqaac cacgggacg tggttttect ttgaaaaca cgataataac 1840 aacgetcaag caccoqaac cacgggacg gascacggtt ttgaaagagg ctccacacg agccggacg taccacggacg ttggggacg cacaagggg gascacgagg gaggaggg tcacaaggg cacaagggg cacaaggggg cacaaggggg cacaaggggg cacaagggg gagggggg tactataggg gagggggg cacaaggggg gagggggg tcacaagggg cacaagggg gagggggggggg	aaccccccac	ctggcgacag	gtgcctctgc	ggccaaaagc	cacgtgtata	agatacacct	T260
taggatette caacaagagg ctgaaagatg cccaagagg acccattgt 1680 ataggatets atetgaggc tegtqaaa tegttaaata (stgattaata gaggttaaa aacgttaaga coccocagaa cacagaggac tegtttaacat (stgattaatac gastaataca aacgttaaga coccocagaa cacagaggac tegtttatacat (stgattaagac cataataca agcctaacaaga cocgagacacaga gaaccaggt gaggggagg tccaagtggt ctcaacaca agcgcaactact tectgacgac ctcaacaga caggacetg cocgctaaaaa cacattacaca accttacaca accattacaca accattacaca accattacaca accattacaca accattacaca agcgacactacacaca agcgacactacacacacacacacacacacacacacacaca	dessaddedd	cacaacccca	ataccacatt.	ataaattaaa	tagttgtgga	aagagtcaaa	1620
ateggatetg atetggggge teggtgage tggtttteet ttgaaaaac ogstaataa 1800 aaogeteag coccogac caegggage tggtttteet ttgaaaaac ogstaataac 1800 agoctoacag geoggaca deceagge agoggggggggggggggggggggggggggggggg	9044499099		30300000	2-2-2-5		200001+64	1600
aatgetctagg coccocqaac cacggggacg tggttttot ttgaaaaaca cgataatacc 1800 atgggcota ttacggcota tcacaaca gacgaggoc tacttgactg catcatacat 1860 agoctcacag gocggacag gaaccaggtc gaggggagg tcacagtggt ctcacacgca 1860 agoctcacag cocggacac ctcstgccgg cacacaggc caggggggggggggggggggggg	tggeteteet	caagegrarr	caacaagggg	cigaaggaig	CCCagaaggt	accedactge	
aatgetctagg coccocqaac cacggggacg tggttttot ttgaaaaaca cgataatacc 1800 atgggcota ttacggcota tcacaaca gacgaggoc tacttgactg catcatacat 1860 agoctcacag gocggacag gaaccaggtc gaggggagg tcacagtggt ctcacacgca 1860 agoctcacag cocggacac ctcstgccgg cacacaggc caggggggggggggggggggggg	atgggatctg	atctqqqqcc	tcqqtqcaca	tgctttacat	gtgtttagtc	gaggttaaaa	1740
atgagagacta ttacagacta deceaaca acgagagac tacttagacta datoacact agoccaaag gocgagaaa gaacaagatcagat gagagagaag tocaaatgag tocaacagacga caacagatcagacaacagatcaga acattacact tectagagac caatacacta gagagagaaga caatacacta tectagagac caaaagaga caatacacca aaatgagac caatgagactag agogactaga agogacacac gagagagacta actateacacc agoccatca tectacacca gagagactaga actatacta gagacactag gagacagac actateacacca gagagactaga catatacacca atgagacacacagagagacacacacacagagagacacacac		2000	as addadas ad	taattttaat	ttgaaaaga	coataatacc	1800
agoctocacag geoggacag gaaccaggtc gagoggagg tocaagtggt ctocacegoa cacacactott tectogogac degetocach ggoctyctyct ggactyctac teatgggcc 1989 ggotcaaaga cecttycogg cocaaagggc ccaatcaccc aaatgtacac caatgtagac 2040 ggogacactog teggtogaca agogcoccoc ggggocgtt cettagacac atgoacctgc ggoggotgacacac ggotctycoc degetogacac tyctotycoc ctogggagacactoc tyctotycoc cogggagacactoc cettogacac tyctotycoc cogggagacactoc cettogacac tyctotycoc cogggagacactoc cettogacac tyctotycoc cogggagacactoc cettogacac agogcocacac tyctotycoc cogggagacactoc cettogacac agoccoctac agoccoctac agoccoctac cyctotycoc coggagacactoc cettogacac accocgaggg typocacactocac agoccoctac agogacactocac caactactycoc cotogacacactac agoccoctac agogacacactac cyctocycocacactac agoccoctac agotacocgac aaactacytoc cotoacactac gaccacacgagacactactactactactactactactactactactactac	aacytttagg	Coccocyaac	cacggggacg	cygrecece	Legadadaca	cgacaacaca	
agoctocacag geoggacag gaaccaggtc gagoggagg tocaagtggt ctocacegoa cacacactott tectogogac degetocach ggoctyctyct ggactyctac teatgggcc 1989 ggotcaaaga cecttycogg cocaaagggc ccaatcaccc aaatgtacac caatgtagac 2040 ggogacactog teggtogaca agogcoccoc ggggocgtt cettagacac atgoacctgc ggoggotgacacac ggotctycoc degetogacac tyctotycoc ctogggagacactoc tyctotycoc cogggagacactoc cettogacac tyctotycoc cogggagacactoc cettogacac tyctotycoc cogggagacactoc cettogacac tyctotycoc cogggagacactoc cettogacac agogcocacac tyctotycoc cogggagacactoc cettogacac agoccoctac agoccoctac agoccoctac cyctotycoc coggagacactoc cettogacac accocgaggg typocacactocac agoccoctac agogacactocac caactactycoc cotogacacactac agoccoctac agogacacactac cyctocycocacactac agoccoctac agotacocgac aaactacytoc cotoacactac gaccacacgagacactactactactactactactactactactactactac	atggcgccta	ttacggccta	ctcccaacag	acgcgaggcc	tacttggctg	catcatcact	
acacaatett tectggegac etgegteaat ggegtgtgt ggactgtcba caatgggac 2004 ggeteaaaga coettegeg caacaaggge caatgegat cettgacac aatgtgac 2004 ggegactegg acetttactt ggteagagg catgegatg teatteeggt ggegeggggggggggggggg	addctdadad	accadaacaa	gaaccaggtc	aaaaaaaaaa	tccaaqtqqt	ctccaccqca	1920
ggctcaaaga cccttgcggg caaaqggc ccaatcacc aaatgtacac aatgtagac 2010 gagaagctogg aggagactog ccttgaacac atgaacctgg 2160 ggcagacacga gggggagact cattacccca aggacagtct cattacctga gggcaggagg cattacctgaga gagagagaga cattaccgga ggggagagactga tactaccaca aggacagta tactacagaga gagagagagacac tgatcacgaga ggtgagactt tgaaaagag ggtgaacttt gtaaccgtcg attataggg aaccactaat gggtacacac aggacagacacacacagggagacacacacacac	ageocoaoag	3003334043		5-55555-55	~~~~	tantaataaa	1000
caggaceteg teggetgga agegecoco ggggegeget cettagaca atgeacteg 2100 ggcagetegg acctttactt ggteacgagg atgecgatg teggegegegg 2160 accepagggg teggegaget actetecoco ageggtecac tgetetgeo cteggagea atcepagg agegegegegg 2280 accepagggg teggaggg ggtggaattt gtaacggtog acctttaga gggetette gtaccgtog agetettegga agegegegg 2280 accepaggg teggaggg ggtggaattt gtaacggtog accetataga acggecata aggaceta aggacetata ggccaaagg taaggggat aaggagat aggaggat acateggoc aggaggatga acateggoc aggagaaggat aaggtgegg cocaaagga taaggggat acateacac tggtategg accateac ggatatatg accacagggt accateac gtaatcoacc accagggt cocacaagga actacacac accacacac accacacacacacacaca	acacaatctt	teetggegae	etgegteaat	ggcgcgcgcc	ggactgttta	ccacggrace	
gagaetteg teggetagae agogecece gggggegett cettagaeae atgeacetge 2100 ggcgacagea gggggaget actetacete ggeggtecae tgetetgee etceggegae ggtggaatet getateeggegegggggggggg	gactcaaaga	cccttaccaa	cccaaaqqqc	ccaatcaccc	aaatgtacac	caatgtggac	2040
ggcgacgocgg ggggagoct actotcocco aggccggtc cotactcgg gcgccggcgg cotoggggg cotogggggggggg	22 casage	tacaataaa	addadadada	agaggagatt	ccttgacacc	atgracetge	2100
ggogacaçaça gggggagoct actotococ aggcocgtt cotacttgas gggcttotoc 2220 ggoggggggggggggggggggggggggggggggg	caggaccicg	coggorggca	agegeeeee	999909090	Coccegacaco	4090400090	
ggogacaçaça gggggagoct actotococ aggcocgtt cotacttgas gggcttotoc 2220 ggoggggggggggggggggggggggggggggggg	ggcagctcgg	acctttactt	ggtcacgagg	catgccgatg	tcattccggt	gcgccggcgg	2700
gggggtcacac tgctctgccc ctcgggggac gctgtgggac tctttcgggc tgcgtgtgc 2340 cggtccagggg tttgcaaggc gaggactct gdaccgtcg gtctatagga aacaactatg 2340 ggccacataga tatagggcac agggaggac catcacacacacacacacacacacacacacacacaca	aacaacaaca	agaggaaget	acteteccce	aggcccatct	cctacttgaa	gaactcttca	2220
accgaeggagg ttetgaaagge gatgaacttt gtaccggteg agtetatatgg aaccactatatg gatgaecteg tetteacgga caatecgtege ceceategae tagtgaegge aagageacta aggtgeegge tgegtatata gegegatatatg ctaaggeaca tggtategae cettegaea cettegaeagge attetgggggga ceceateagggt gedtatatge ctaaggeaca tggtatecgae ctataggaagg cecaacaca gaaccacagggt accacacacacacacacacacacacacacacacacaca	ggogacagca	5555545000		~355	+a+++aaaaa	tagaatataa	2280
cggtcccagg tettcaccag caactegtee cetceggeg tacegaga attecaggtg 2460 geccatettae aegeceteae tygtagegg aagageacta aggtscegge tygtatagoa 2460 geccaagggt ataaggteet tygtectgaae cgtcatagtge cetcagace taggagege cetatageae tygtategae cgtcatageage cetcagage coactagage cetatagaca taggagegeggegggggggggggggggggggggggggg	ggcggtccac	tgetetgeee	creggggeae	gergraggea	LCLLLCgggc	tgeegegege	
gecteateda accecetae tygtageges aagageacta agytegege tygetatagea cegecaagagt ataaggtget tyteetgaac egytatetge cetaacaaca egytatega cacaggytg ceccaacaggytg ceccaacaggytg ceccaacaca gytatecgac tygtategac cataatatag gatagagge accetatagagt tygetategac cataatatag gatagagge accetatagaaca tygtategac cataaatatag gatagaggeg cectatgacat cataatatag gatagaggeg accetagaca tectgagaca tygetategac cataaatatag gatagagggagga accetagaca tygetategac cataaatatag gatagaggggaga cectagacaca cataatatag gatagagggaggagaggggggggagacacaa gatagagaggggggggagaggaga	acccgaggg	ttqcqaaqqc	ggtggacttt	gtacccgtcg	agtctatggg	aaccactatg	2340
geceatetac acgecetac tggtagege aagageacta aggtgeege tgegtatgea 2460 geceaagggt ataaggtpet tgteetgaac eegteegteg eegeeacet aggttteggg 2520 gegtatatge 2520 gegtatatge 2520 gegtatatge 2520 gegtatatge 2520 gegtatatge 2520 getageaggg eegeeacat etggtategga eegeacatge 2520 getetggege eegeacatge eetggaecatge 2520 getetgeege geceaatge eetggaecatge 2520 getetgeege geactggee eeggaecatge 2520 getetgeege geactggee eeggaecatge 2700 aagggaggg gegaecetat tteetgeeaat eatggeaag eatecacetg 2700 aagggaggg gegaecetat tteetgeeat teatggeaag eatecacet 2920 getetgeege geetgggag aaggeecate 2700 aagggagggg gegaecetat tteetgeeat teatggeaag eatecaceat 2920 gegaecatg 2820 getetgeege geetggga aatgetgte 2700 aagggagggg gegaecetat tteetgeeat teatggeaag eatecaceat 2920 gegaecate 2880 aaggggggg gegaecetat tteetgeeat teatggeaag eetcaaceatg 2920 gegaetgte 2920 geatattaee 2920 geatattaee 2920 gegaetteg 2920 geatattaee 2920 geegaace 2920 geegaacet 2920 geegaacet 2920 geegaacet 2920 geegaacet 2920 geegaacet 2920 geatattaee 2920 geegaacet 29	aaataaaaa	tattazaaa	caactcotco	cet ceaacea	taccocadac	attccadata	2400
goccaagggt ataaggtgot tytootgaac cotacoaca gaaccacet aggtttoggg goctataty otaaggoaca tygtatogac cotacacac gaaccacgggt accaccacac gaaccacgaggt accaccacac agaaccacac accacgggt coccatcac gaaccacac tatgcaagt tottpaggagg cotatgacat cataataty gatgagtgoc actoacacac catagacacac catagaggagga cotatgacat cataataty gatgagtgoc actoacacac catagaggagga cotacgacac gagaacgag catggagaggg cotaggacacac gagaacgag catggagaggg gaccaccacat ttottgocat tocaagaaga aatgtgatga gagaacgtc catggagacaa gaagggggga gocacctoat tttotgocat tocaagaaga aatgtgatga gagaacgtg cattggta gaaacggaggg cattgacacacacacacacacacacacacacacacacaca	eggedeedgg	correacyga	caactegeee	cccccggccg	caccycagac	h	
gcgtatatgt chaaggcaca tggtatcgac ccgtacgtcg ccgcaccct aggttteggg gcgtatatgt chaaggcaca tggtatcgac cctaacatca gaaccggggt aggaccatc 2580 accacgggtg cccccatcac gtactccacc tatggcaagt ttettgccga cggtggttgc 2640 accacgggtg ccccacatca cataatatgt gatgacgac atctaactga ctcgaccact 2700 atcctggacgc cctagcacca cataatatgt gatgacgac atctaactga ctcgaccact 2700 accacggggggggac tcggcacact cataatatgt gatgacgac atctaactga ctcgaccact 2700 accacgacggggggggacacggcaccact 2700 accacgacgacgacaccaccact 2700 accacgacgacgacaccaccact 2700 accacgacgacgacacaccaccactactgacgacacaccaccacaccacacaca	gcccatctac	acgcccctac	tggtagcggc	aagagcacta	aggtgccggc	tgcgtatgca	2460
gogtatatatyt ocacegoaca tygtatecgac octaacataa gaacoggggt aaggaccato 2580 accacgggtg occceatoac gtactocace tatggcaagt tecttgogga octatgacat cataatatyt gatgagtgcc actocacct 2700 atcotggggc teggacaag catgagacga ctoggacacat 2700 atcotggaga teggacaagt catgagacga ctoggacacat 2700 ctogcaccc deagcactc gggategytc accetgacac catacacata catgagaaga catcacacat 2920 gctotytoca gcactggaga aatocacettt tatggcaaag ccatcocaat cgagacactc 2880 aaggggggga ggcactcat tttotgccat tocaagaaga aatytgatga gctocgccgg actcytocg gcotcygaccac caatgctyta gcaacagaaga aatytgatga ggcactcacacact geggagacgt cattyctaa gcaacagaaga aatytgatga ggcactcacacact gagaccac acacacacacacacacacacacacacacac	acccaagggt	ataaggtgct	totcctgaac	ccatccatca	ccaccaccct	aggtttcggg	2520
accacgggtg ceccatea gtactecace tatggeaagt ttettgeega eggtgttge 2640 atcagggeg ctatgacat cataatatgt gatgagtgee acteaactga atcageagt teggaegaeg ctagagegaeg etggaegaeg actegaegaeg actegaegaeg actegaegaeg actegaegaeg actegaegaeg actegaegaeg actegaegaeg actegaegaeg eggagaegg etggaegaeg actegaegaeg actegaegaegaegaegaegaegaegaegaegaegaegaegae	5000445550		+ catataca	aataaaataa	as sacacat	aaddaddatd	2580
tetggaggeg cetatgacat cataatatg gatgagtgec acteaactga ctegacact tetggacaa geggagaggg ctggagggg catgacggt categacat cetggacaa geggagagggg ctggaggggg aatceaacat gaggagggg aatceactt tatggaaaag catcaacat gagagagggg gacgetgat tettetgeat tecaagaaga aatgatgatga gagaggggggggggggggg	gegtatatgt	ctaaggcaca	tygtategae	CCLaacatca	gaaccggggt	aayyaccarc	
atctgggggg ctatgacat cataatatgt gatgagtgc actcaactga ctcgaccact 27700 atctgggggga teggcacagt cetggaccaa geggagacgg ctggagaggg actcgtcgtg 2760 ctcgccaccg ctacgcetce gggateggte accgtgcac atccaaacat egagagggg geactgtaca aggaggggga aatcccctt tatggcaaag cattgatga gatgcgggggggggg	accacqqqtq	ccccatcac	gtactccacc	tatggcaagt	ttcttgccga	cggtggttgc	2640
atoctgggca toggcacaag ctagaccaa goggaagacg ctggaagacg actcgtctgd 2760 ctcgccaccc ctacgcctcc goggatcggc acceptacea acceptacea cacaacaac cacaacaac acceptacea aaggagggga aatocccttt tataggcaaag catccaacacat cgaggaacatc 2880 aaggagggga gocactcat ttctgcaat tocaagaaga aatotgccgg 2940 aaggatggc catcggaacg catcgtgta goaacggacg ctctaatgac gggcetttacc accaacacta goggaacgac catcgtgta gaacaggacg ctctaatgac gggcetttacc cggagacgg cattggta gaacaggacg ctctaatgac gggctttacc cggagacgg catcgaacgac cacaacgac ccaacacacacacacacacac	+a+aaaaaaa	cctatcacat	cataatatot	dataadtacc	actcaactga	ctccaccact	2700
ctogcacog ctacgoctoc gggatoggt account tatggocaca atoccaatact cgagagagtg ggcactotytoca goactoggaa aatoccottt tatggocaaag ccatoccocat ggagaccato aagggggggg ggcactoct tttctgocat toccaagaaga aatgtgatga gctoggocg 2940 aaggtgtocg gcotoggact caatgotyta goatattaco ggggotttac ggcgatttog actoagtgat cgattgotya goaacggacg cottoaccat taggacgacg cottoaccat taggacgacg acgggocact caggacgacg cottoaccat taggacgacg acgggocac aaggggogg gcaggocgg cottoaccat taggacgacg acgggocac aaggggogg gcaggocgg acgggocac caggggocac cagggocac cagggocacg agggcatyt acagggocac aaggggocacg aggcatyt acagggocac acgggocacg aggcatyt acagggocac acgggocacg aggcatyt acagggocac acgggocacg aggcactyt acagggocacg acgggocacg aggcactyt acagggocacg aggcactyt acagggocacg acgggocacg aggcactyt acagggocacg aggcactyt acagggocacg aggcactyt acagggocacg aggcactty gccgagacc caccacacac aggcocacty ggaccactory gtctgtogg agtgotaty ggcggoctta accacacac acccacatag acgcccattt cttgtoccag ggattotogg aggaggacac cttcacacac acccacatag acgcccattt cttgtoccag actaagoag actacacgac acccacacac acccacacac acccacacaca	55599999	'		33-3-3-	atacaracaa	nataataata	
gotcetyteca goactogaga aatecettt tataggeaaag ceateceat egagaeeate aaggegggga goaceteat ttetgeat teeaagaaga aatgtgatag getegeegeg 2940 aaggegtteg geeteggaet caatgetgta geaacggaeg etetaatgae ggggetttaee gggggetttaee egageggeggattteg acteagtgat egactgeaat acatgtgtea ecagaeage etetaatgae gggetttaee egageggegggggggggg	arccrgggca	teggeaeagt	cetggaecaa	geggagaegg	ctggagegeg	actiguig	
getetgteca geactggaga aatececttt tatggeaaag ceateceat cgagaceate 2880 aaggggggag ggeaceteat ttetegeeat tecaagaaga aatgtgatga getegeegeg 2940 aagetgteeg geeteggaet eattgtegta geatattaee ggggeettga tgtateegt 3000 ataceaacta geggagaegt cattgtegta geatattaee ggggeetttaee cgageagaeg etetgaeaet acatgtgtee eccagaeagt eggetttaee cgageagaeg eeggagaegg etetgageegga geaggaetgg taggaggaegg acceptgeae aagaeggegt geagggeegg gageagaegg eggeggaattt acaggtttg gactecagga 3120 ggacggeggaggaggaggaggaggaggaggaggaggaggaggag	ctcqccaccq	ctacqcctcc	gggatcggtc	accgtgccac	atccaaacat	cgaggaggtg	2820
aagggggggg ggcacctcat tttotgcat tccaagaaga aatgtgatga gdtcgccgg ggaagactgaattcgatgtgta gcaatgtgta gcaatgtgta ggggcttga tgtatccgtc 3000 ataccaacta goggaacgt cattgtcgta gcaacggacg ctctaatgac ggggctttac catgatgat caccaacta goggaactg cattgtcgta gcaacggacg ctctaatgac gggctttacc caggacgcatttcg actoagtgat cgactgcaat acatgtgtca cccagacagt cagggccacga ccttcaccat tgagacgacg accgtgcac aagacgcggg ggcactcaggacggacggacggaaggacggacggacggac	actitatica	acactagaga	aatccccttt	tatoocaaao	ccatccccat	cgagaccatc	2880
aagetyteeg geeteggaet ataecaacta geggaaget edtygaeega ettegea ettegegae edtygaeega ettegeaegaegaegaegaegaegaegaegaegaegaegaega	5000050000	3000033030	tttataaaat	+442242243	aatataataa	acteaceaca	2940
ataccaacta geggagacyt cattycgta geaacggacy ctctaatgac gggetttace cagedyatteg acteacytgat cagacyteat acatytytea cetteaceat tgagacyacy accyteacya accyteacyacy gaagacytey gaagacyacy cagedyatyt cagacygeay atgygeatt acagytytyt gactecagya 3240 acagygtyty cagacyacy cagacyted acagygtyty cagacyacy cagagyacy acgygeayth acagytytyt gactecagya 3240 acagygtyty cagacyacy cygeatett cagagytyty agygeayth acagygytyty agygeayth acagygytyty agygeayth acagygytyty cygeatett acagygetty agygeatt acagygytyty agygeayth acagygytyty acygaayytytytytytytytytytytytytytytytytyty	aaggggggga	ggcaccical	LLLCLGCCaL	cccaagaaga	aacycyacya	geregeege	
ataccaacta geggagacyt cattycgta geaacggacy ctctaatgac gggetttace cagedyatteg acteacytgat cagacyteat acatytytea cetteaceat tgagacyacy accyteacya accyteacyacy gaagacytey gaagacyacy cagedyatyt cagacygeay atgygeatt acagytytyt gactecagya 3240 acagygtyty cagacyacy cagacyted acagygtyty cagacyacy cagagyacy acgygeayth acagytytyt gactecagya 3240 acagygtyty cagacyacy cygeatett cagagytyty agygeayth acagygytyty agygeayth acagygytyty agygeayth acagygytyty cygeatett acagygetty agygeatt acagygytyty agygeayth acagygytyty acygaayytytytytytytytytytytytytytytytytyty	aagctgtccg	gcctcggact	caatgctgta	gcatattacc	ggggccttga	tgtatccgtc	3000
ggggatttcg cttgaccaga cattaggacga acastgsta cacagacaga cagactcagc cuttcacacat tgagacgacga accgstgcaa aagacgacgg ggacagactgg tagggcagag acgggcatta acaggststst gagacgacgag atgggcatta acaggststst gagacgacgag atgggcatta acaggststst gagacgacac aggacatgst caggscagagacggatgst cagstcatgagagacgagacgagacgagacgagacgagac	ataccaacta	acaaaaacat.	cattotcota	acaacaaca	ctctaatgac	gggctttacc	3060
cagagaccaga cetteacat tgagacgacg accgtgocac aagacgcggt gtcacgctcg algagactgg cagagactgg tagggcaag atgggcatt acaggtttgt gactecagga 3240 gaacgaccet cgggcatgtt cgattecterg gtcttgtgcg agtgetatga cgcgggctgt cgcagagacc tcagttaggt tgcggggtta cctaacaca accacatag acgccattt ctgtcacag accacactag accacactagagaggcaggactagag accacactagagaggcaggaggactggt acacacacac accacactagagaggcaggactggaggactggaggactggaggatgggactgagaggactggaggaggaggaggaggaggaggaggaggaggaggagga		5055454050		5555	aaaaaaaa+	adaattaada	3120
cagegagag geagactgg tagggeagg atgggeattt acaggtttgt gactcagga gatggeagg geatggeagg geatggeagg gettegeagg gettegeagg gettegeagg gegggettt ceagttaggt tgggggettt acaggeagg ceagggttg cegttegeagg ggacatetg gagttetggg agggettt tacaggeete accacatag cegecattt cttgteceag actaageagg caggagacaa cttecectac accacatag cegecattt cttgteceag actaageagg caggagacaa cttecectac accacatage cegetteaga accaeggetac acaatgtgga agtgteteat acaggetaca accacaca acggetaca acggetaaag cetaageagg ctcaaggetacaggeteetatagggatggatggatggatggatggatggat	ggcgattteg	acceagigai	cyactycaat	acatytytea	cccagacage	cgaccccagc	
cageggegag gaaeggeete cgggattyt gettggtacg gettggtacg gettggtacg gettggtacg ccaeggttg ccaeggttg ccaeggttg ccaeggttg ccaeggttg ccaeggttg ccaegagac ccaeatag ccaeatag ctggtagaa ccaeatag ctggtagaa ccaeatag ctggtagaa ccaeatag ctggtagaa caaattgtgga agtgtetcaa agggtataaag gatgetcagg atcaeateg agggetagg agggetagg cttggtagaa ccaaatgtgga agggetaca caaatgtgga agggetaca caaatgtgga agggetaca caaatgtgga agggetaca agggetaaag cctaeggtg caeaateca agggetaaag cctaeggaga ccaegagaca cctaeggaga cctggggaga cctgggaga cctggggaga cctgggaga cctgggagaga cctgggaga cctgggaga cctgggaga cctgggaga cctgggaga cctgggaga cctgggagaga cctgggagaga cctgggagaga cctgggagaga cctgggagaga cctggagagaga cctaeacac ccaeacaga ccaeacaca ccaeacac	ctggacccga	ccttcaccat	tgagacgacg	accgtgccac	aagacgcggt	gtcacgctcg	3780
gaacggccct cgggcatgtt cgattcctcg gttctgtgcg agtgctatga cgcgggctgt 3300 cctaggtacg agctcacgcc cgcgagacc tcagttaggt tgcgggctta cctaaacaca 3360 ccagggttgc ccgtctgcca ggaccatctg gagttctggg agaggggcgctt tacaggcctc tacaggagaca cttccccacacacacacacacacacacacacacaca	caucaacaaa	acadaactaa	tagggggagg	atoggcattt	acaggtttgt	gactccagga	3240
gettgstacg ageteaegee egeegagace teagttaggt tgegggetta cetaacaca 3360 cagggttge cegtetgeea ggaceatetg gagttetggg agaggggetet tacaggeete 3420 caaatgtgga agggeteat accaggetae ggtgtgegee agggeteagg etecaetee ategtgggac 3540 acggeatagg etetaggetga agggeteate acaggeegetaaag eteaegggeegetaagg eteaegggeegetaagg eteaegggeege agggeegeegetgggaeggeegeteagg eteaegggeegeegeegggeegeegggaeggeeggeegge	cagoggogag	3043340433				acacacatat	
ccagggttgc accacatag accacatct cttgtcccag actaagcagg caggagacaa cttcccctac accacatag accacatag accacatct cttgtcccag actaagcagg caggagacaa cttcccctac accacatag aggtctctat cttgtcccag actaagcagg caggagacaa cttcccctac atcgtggac agggctcagg ctccacctcc atcgtggac actaagcagg ccaaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gccctgtg actacgatgca attaaggctgg aggccgttca acaacagaggt actaccacac acccataaac acccataaacat gtgcatgcag ctctgggggg gtattggcgg gtattgcggg accatcatt cccgacaggg aagtccttta cgggaggtc gtgaacaattca acacaacaca	gaacggccct	egggeatget	egatteeteg	gricingly	agigolatga	cacadaccac	
ccaggttge ccgtetgea ggaccatetg gagttetggg agagggtett tacaggete accacatag acgecattt ettgteecag actaaggagg caggagacaa ettggtggac acgaggagaca accacatag acgaggteagg ettgggac aggggetagg ettgggac attgggac acggggetaagg etcacacetee acggetaaag ectacgetge acggecaaca acgecatgagac atggecatgaggt etcacagagaca etcatagagagaca etcatagagagaca etcatagagagaca etcatagagagaca etcatagagagaca etcatagagagaca ectacagagaca acaacaaggaa acaacaaggaa acaacaaggaa acaacaaggaa acaacaaggaa eccacaaggaagagagagagagagagagagagagagagag	gcttggtacg	agctcacgcc	cgccgagacc	tcagttaggt	tgcgggctta	cctaaacaca	3360
acccacatag acgcccattt cttgtcccag actaagcagg caggagacaa cttccctac 3480 ctggtagacat accaggctac ggtgtgcgcc agggctcagg ctccacctcc atcgtgggac accaatgtgga agtgtctcat acggctaaag cctacgctgc acggccaac gcccctgctg 3600 atggcatgca tgtcggctga cctggaggtc gtcacaggca accccataac caaatacatc 3660 atggcatgca tgtcggcgc gtattgcctg accacacac accccataac caaatacatc 3660 atggcatgca tgtcggcgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg 3720 gtcctagcag ctctggcgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg 3780 atcatcttgt ccggaaagcc gccatcatt cccgacaggg accacattca accagaaggc ctcacacctc ccttacatcg aacagggaat gcagctcgcc gaacaattca accagaaggc accetcacacac accacaagca agcggggggattca acaatattta gcaggcttgt ccaccacaca accccaacaca agcggagggg gcgaattcat tcagcggggat accatattta gcaggcttgt ccaccacacac accccaacat 4020 gcgatagcat cactgatggc attcacacgc tctatcacac gcccgctcac cacccaacat 4140 acctcctgt tggatatttt ggaggggggggggggggg	casaaattaa	ccatataca	agaccateta	gagttctggg	agagggtgtt	tacacccccc	3420
ctggtagcat accaggetac ggtgtgegc aggggttagg ctcacactcc atcgtggac 3540 caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gccctgctg 3600 tataggctgg gagccgttca acacagaggtt actaccacac accccataac acgccatgca gtcatagcag ctctaggaggtc gtcacgagga cctgggtgct ggtaggcgga 3720 gtcctagcag ctctggaggtc gtcacgagga cctgggtgct ggtaggcgga 3720 gtcctagcag ctctggaggtc gtcacgagga gcgtggtcat tgtgggcagg 3780 atcatcttgt ccggaaagcc ggccatcatt cccgacaggg acacactca accacacac accacaggaat ggatgaggagga acacactca accacacac accacaagga accacacaggaat gcagcacacacacacacacacacacacacacacacacac	ccagggccgc	cogeologica	ggaccaccag	9490000999			
caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gcccctgctg gagccgttca aaacgaggtt actacgctgc gagccgttca aaacgaggtt actacgacac accccataac caaatacatc 3660 atggcatgca tgtcggcgc gtattgcctg gtcacgagca cctgggtgct ggtaggcgga 3720 gtcctagcag ctctggcgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg 3780 atcatcttgt ccggaaagcc gccatcatt cccgacaggg aagtccttta ccgggagttc gatgagatgg aagagtgcgc ctcacacctc ccttacatcg acacgggaat gcagctcgcc gacacaattca aacagaaggc atcgggtgg accetcgaag ccttctgggc gaagattcgc caagtggcgg accetcgaag ccttctgggc gaagatattg tggaaattca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaaccc 4080 gcgatagcat acactgatggc acggagggtt tgggagggct tgggagggct tgggagggct ttggcaggagggct tgggagggct tgggagggggggggg	acccacatag	acgcccattt	cttgtcccag	actaagcagg	caggagacaa	Culdedelad	
caaatgtgga agtgteteat aeggetaaag cetaegetge aegggeeaae geeetgetg 3600 ataaggetgg gageegttea aaacgaggtt actaecacae accecataae caaatacate 3660 ataggeatgea tgteeggege gtattgeetg gteaegagea gegtggtest tgtgggeagg 3720 gteetageag etetetgge gtattgeetg acaacaggea gegtggteat tgtgggeagg 3780 ateatettgt eeggaaagee gecateatt eeggacaggg aagteettta eeggaggte gaacaattea aacagaagge etecaeacee eettacateg aaacagggaat geaggetegee gaacaattea aacagaagge etecaeacee eettacateg eeaceeaggaattee geaggetegee ggaacaattea teagegggat eeaceteegaag eectetegge gaaggattgt geagattee eeaceteetge geagatateg eeaceteetge geagatateg eeaceteetgee geagatateg eeaceteetgee geagatateg eeaceteetgee geagatateg eeaceteetgee teesaaggeetgeet teesaaggeetgeetgeetgeetgeetgeetgeetgeetgeetg	ctqqtaqcat	accaggctac	ggtgtgcgcc	agggctcagg	ctccacctcc	atcgtgggac	3540
tataggetgg gagecgttca aaacgaggt actaccaca accccataac caaatacatc atggcatgca tgtcggcgg gtattgcctg gtcacgagca cctgggtgct ggtaggcgga 3720 gtcctagcag ctctggcgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg 3780 atcatcttgt ccggaaagcc gccatcatt cccgacaggg aagtccttta ccgggagttc gatagagatgg aagagtgcgc ctcacacctc ccttacatcg aacagggaat gcagctcgcc 3900 gaacaattca aacagaaggc aatcgggttg ctgcaaacag ccaccaagca agcggaggct ggaaattca tcacgggat caagtggcgg accctcgaag ccttctgggc gaagcatatg 4020 tggaaattca tcacgcggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080 gcgatagcat cactgatggc attcacaccac ggcggcgcc aacttgctcc tcccaagcac 4200 gcttctgctt tcgtaggcgc cggcatcgct ggaggcggct ttggcaggcg gctcgtggc gatggcctcac accgaggac tggcggcgc gatggcgcc acctctgcc tggcagccc 4320 aaggtgcttg tggatatttt ggcaggttat ggaggaggg tggcagggcg ggtggagggc tggtgggcgcac agctggcg accctcgga acccctgga acccctgga acccctgga agccttggc ggagggcc tggtgggccc acctggc acccctactgc ttggcaaccc 4380 atcctctccc ctggcgccc agcctatgtg ccacacacac tggtggcccac agccgatact gggaggggc tggtggggacg atgaccccc accgaggac tggtggcgcac acctcggg accctcggg accctcc accgaggac tggtggccac acctcgcg 4500 ggtaaccacg tctccccacac gcactatgtg ccacacacac ggggagacgg accctcacacacacacacacacacacacacacacacaca	casatritros	actototoat	accortaaac	cctacactac	accoccaac	acccctacta	3600
atggcatgca tgtcggctga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 3720 gtcctagcag ctctggccg gtattgcctg acaacaggca gggtggtcat tgtgggcagg 3780 atcatcttgt ccggaaagcc ggccatcatt cccgacaggg aagtccttta ccggaggttc gacaacattca aacagaaggc accttacatcg accacacacacacacacacacacacacacacacacac	cadacgegga	agogoooao				ganatagata	3660
atcatcttyt ceggaaagee geatetett ceegacaggg aagteettta eegggagtte 3840 gatgagatgg aagatgege etcacacete cettacateg aacaagggaat geagetegee 3900 gaacaattea aacagaagge aategggty etgecaaacag ceaccaagga ageggagget 3960 getgeteeeg teggtggaate caagtggegg accetegaag cettetggge gaageataty 4020 teggaattea teageggat acaatatta geaggettyt eeactetgee teggaaacee 4080 gettetget tegaaggege eggeateget gaagggget teggagget teggaggget teggaggget teggaggget teggagggggggggg	tataggctgg	gageegeeea	aaacgaggtt	actaccacac	accccataac	Caaacacacc	
atcatcttyt ceggaaagee geatetett ceegacaggg aagteettta eegggagtte 3840 gatgagatgg aagatgege etcacacete cettacateg aacaagggaat geagetegee 3900 gaacaattea aacagaagge aategggty etgecaaacag ceaccaagga ageggagget 3960 getgeteeeg teggtggaate caagtggegg accetegaag cettetggge gaageataty 4020 teggaattea teageggat acaatatta geaggettyt eeactetgee teggaaacee 4080 gettetget tegaaggege eggeateget gaagggget teggagget teggaggget teggaggget teggaggget teggagggggggggg	atggcatgca	tqtcqqctqa	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	3720
atcatcttgt ccggaaagcc ggcatcatt cccgacaggg aagtccttta ccgggagttc 3840 gatgagatgg aagagtgcgc ctcacacctc ccttacatcg aacagggaat gcagctcgcc 3900 gaacaattca aacagaaggc aatcggggg accetcgaag ccttctgggg gaagcatatg 4020 tggaattca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaacccc gcgatagcat cactgatggc atcacagcc tctatcacca gcccgctcac caccaagca 4140 accetcctgt ttaacatcct ggggggatgg gtggccgcc aacttgctcc tccaagcgct tctacgct tcgaagcgct tcgaagcgct ttggcagcat aggccttggg aaggtgtg tggaagcatatt ggagggggtg tggaagcatatt ggagggggggggg	atacteacea	ctctaaccac	gtattgcctg	adaadaggda	acataatcat	tataaacaaa	3780
gatgagatgg aagagtggg ctcacacctc ccttacatcg aacagggaat gcagctcgcc 3900 gaacaattca aacagaaggc aatcgggtg ctgcaaacag ccaccaagca agcggaggct 3960 gctgctcccg tggtggaatc caagtgggg accetcgaag ccttctgggc gaagcatatg 4020 tggaatttca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080 gcgatagcat cactgatggc attcacagcc tctatcacca gcccgctcac cacccaacat 4140 accetcctgt ttaacatcct ggggggatgg gtggccgcc aacttgctc tccaagcgct tcgtaggcg cggcatcgct ggagggggtg ttggcagcac aggccttggg 4260 aaggtgcttg tggatatttt ggcaggttat ggagcagggg tggcaggcg gctcgtggcc 4320 tttaaggtca tgagcggca gatgcctcc accgaggacc tggtaacct actccctgct 4380 atcctctccc ctggcgccc agtcgtgg atgaccacg tggtaaccac gggaggggc tgtgcagtgg atgaccgcc tgatagcgt tggtagcgc 4440 gtgggcccag gggaggggc tgtgcagtgg atgaccggc acgctgcacc accgaggac 4500 ggtaaccacg tctccccac caccacac ctgctgaaga ggcttcacca gtggatcact 4560 aagatcctc ctagtctac catcactcag ctgctgaaga ggcttcacca gtggatcact 4680 agggactgct ccacgccatg ctccggctcg tggctaagaa atgtttggg atggatcaac 4680 agggactgct ccacgccatg ctcccggctcg tggctaagaa atgtttggg atggatcacc 4680	5000005				andtagttta	acacacatta	3840
gaacaattca aacagaaggc aatcgggtg ctgcaaacag ccaccaagca agcggaggct 3960 gctgctcccg tggtggaatc caagtgggg accetcgaag ccttctggg gaagcatatg 4020 tggaattca tcagcgggat acaatatta gcaggcttgt ccactctgcc tggcaacccc 4080 gcgatagcat cactgatggc atcaccagcc tctatcacca gcccgctcac cacccaacat 4140 accetcctgt ttaacatcct ggggggatgg gtggccgcc aacttgctc tccaagcgct tcgtaggcgc cggcatcgct ggagggggtg ttggcagcat aggccttggg 4260 aaggtgcttg tggatatttt ggcaggttat ggagcagggg tggcaggcg gctcgtggcc 4320 tttaaggtca tgagcggca gatgccctcc accgaggacc tggtaacct actccctgct 4380 atcctctccc ctggcgccct agtcgtcggg gtggtggg tggtaggcgc tggtaacct gcgtcggg 4320 ggtaaccacg tctccccac gcactatgtg cctgaggcg tgatagcgt tggtaacct gggtaggcc 4440 gggaaccacg tctccccac gcactatgtg cctgagagcg acgctgcacc acgtgtcact 4560 aggactcctc ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcacc 4680 aggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttgggatcacc 4680 aggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatcacc 4680	atcatcttgt	ccggaaagcc	ggccaccacc	cccgacaggg	aagteetta	ccgggagttc	
gaacaattca aacagaaggc aatcgggttg ctgcaaacag ccaccaagca agcggaggct 3960 gctgctcccg tggtggaatc caagtggcgg accetcgaag ccttctgggc gaagcatatg 4020 tggaatttca tcagcggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080 gcgatagcat cactgatggc attcacca gccgctcac gccgctcac cacccaacat ttaacatcct ggggggatgg gtggccgcc aacttgctcc tcccaagcgt 4200 gcttctgctt tcgtaggcgc cggcatcgct ggaggcgggt ttggcaggcac aggccttggg 4260 aaggtgcttg tggatatttt ggcaggttat ggaggagggg tggcagggg tggcaggcg gctcgtggcc atcctctccc ctggcgcca agtcgctcc accgaggacc tggtaacct gcgggagggg ggtgggggggggg	gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacagggaat	gcagctcgcc	
gctgctcccg tggtggaatc caagtggcgg accetcgaag cettctgggc gaagcatatg tggaatttca tcagcggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080 gcgatagcat cactgatgcc attcacagcc tctatcacca gcccgctcac cacccaacat tccactcctgt ttaacatcct ggggggatgg gtggccgcc aacttgctcc tccaagcgct tctaaggtcat tcgtaggcgc cggcatcgct ggagggggtg ttggcaggcat aggccttggg 4200 aaggtgcttg tggatatttt ggcaggttat ggagggggggggg	caacaattca	aacagaaggc	aatcoootto	ctocaaacao	ccaccaaqca	agcggaggct	3960
tggaatttca tcagcggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080 gcgatagcat cactgatggc attcacagcc tctatcacca gccgctcac cacccaacat 4140 accctcctgt ttaacatcct ggggggatgg gtggccgcc aacttgctc tccagcgct tcgtaggcgc cggcatcgct ggagcggctg ttggcagcat aggccttggg 4260 aaggtgcttg tggatatttt ggcaggttat ggagcagggg tggcagggc gctcgtggcc 4320 tttaaggtca tgagcggca gatgccctcc accgaggacc tggtaacct actccctgct 4380 atcctctccc ctggcgccct agtcgtcgg gtcgtgtgcg cagcgatact gcgtcggg 4320 gggagggccag gggaggggc tgtgcagtgg atgaccgcc tgatagcgt cgcttcgcgg 4500 ggtaaccacg tctccccac gcactatgtg cctgagagcg acgctgcacc accgtgcac 4500 aggaaccgct ctagtctacc catcactcag ctgctgaaga ggcttcacca gtggatcact 4560 aggactgct ccacgccatg ctccgggct tggctaagag acgctgcacc acgtgatcacc 4680	gaacaaccca	L			aattataaaa	annage teta	4020
gcgatagcat cactgatgc attcacagc tctatcacca gcccgctcac cacccaacat 4140 accctcctgt ttaacatcct ggggggatgg gtggccgcc aacttgctc tcccagcgct 4200 gcttctgctt tcgtaggcg cggcatcgct ggagcggctg ttggcagcat aggccttggg 4260 aaggtgcttg tggatatttt ggcaggttat ggagcagggg tggcaggcg gctcgtggcc 4320 tttaaggtca tgagcggca gatgccctcc accgaggacc tggtaacct actccctgct 4380 atcctctccc ctggcgcct agtcgtcggg gtcgtgtgcg cagcgatact gcgtcggcac 4440 gtgggcccag gggaggggc tgtgcagtgg atgaaccggc tgatagcgt cgcttcgcgg 4500 agtaaccacg tctccccac gcactatgtg cctgagagcg acgctgcac acgtgtcact 4560 aggactctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcacc 4620 agggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatcacc 4680	getgeteeeg	rggrggaard	caaguggugg	accelegaag	CCEECLGGGC	gaagcatatg	
gcgatagcat cactgatggc attcacagc tctatcacca gcccgctcac cacccaacat 4200 accctcctgt ttaacatcct ggggggatgg gtggccgcc aacttgctcc tcccagcgct 4200 gcttctgctt tcgtaggcgc cggcatcgct ggagcggctg ttggcaggcat aggccttggg 4260 aaggtgcttg tggatatttt ggcaggttat ggagcagggg tggcaggcgc gctcgtggcc 4320 tttaaggtca tgagcggca gatgccctcc accgaggacc tggtaacct gcgtcggcc atccctccc ctggcgccct agtcgtcggg gtcgtgtgcg cagcgatact gcgtcggcac 4380 ggtaggcccag gggaggggc tgtgcagtgg atgaaccagg tgatagcgt cgcttcgcgg atgaaccacg tctccccac gcactatgtg cctgagagcg acgctgcacc accgtgcacc 4500 ggtaaccacg tctccccac gcactatgtg cctgagagcg acgctgcacc acgtgtcact 4560 agggactgct ccacgccatg ctccggctcg tggctaagag acgttcacca gtggatcacc 4620 agggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatcacc 4680	tggaatttca	tcagcgggat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
accetcetgt ttaacateet ggggggatgg gtggeegeec aacttgetee teccageget 4200 gettetgett tegtaggege eggeateget ggageggetg ttggeagged aggeettggg 4260 aaggtgettg tggatatttt ggeaggttat ggageagggg tggeaggege getegtggee 4320 tttaaggtea tgageggea gatgeeetee accegaggaee tggtaacet acteeetget 4380 atceteteee etggegeeet agtegteggg gtegtgtgeg eagegataet gegteggeae 4440 gtgggeeeag gggagggge tgtgeagtgg atgaacegge tgatagegtt egettegeg 4500 ggtaaceaeg teteceeeae geaetatgtg eetgagageg acgetgeaee 4500 aggtaaceaeg teteceeeae etagtettae eateaeteag etgetgaaga ggetteaeea gtggateaee 4620 agggaetget eeaegeeatg eteeggeteg tggetaagaa atgtttggga ttggateaee 4680	acastsacst	cactgatggc	attracador	tctatcacca	geeegeteae	cacccaacat	4140
gcttctgctt tcgtaggcgc cggcatcgct ggagcggctg ttggcagcat aggccttggg 4260 aaggtgcttg tggatatttt ggcaggttat ggagcagggg tggcaggcgc gctcgtggcc 4320 tttaaggtca tgagcggcga gatgccctcc accgaggacc tggttaacct actccctgct 4380 atcctctccc ctggcgccct agtcgtcggg gtcgtgtgcg cagcgatact gcgtcggcac 4440 gtgggcccag gggaggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgcgg 4500 ggtaaccacg tctcccccac gcactatgtg cctgagagcg acgctgcagc acgtgtcact 4560 cagatcctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac 4620 gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatcatc 4680	Jogaragear					taggaggggt	
aaggtgettg tggatatttt ggcaggttat ggagcagggg tggcaggcgc getegtggee 4320 tttaaggtea tgageggega gatgeeetee acegaggace tggtaacet actecetget atceteteee etggegeeet agtegteggg gtegtgtgeg eagegatace gegteggeae 4440 gtgggeeeag gggagggge tgtgeagtgg atgaacegge tgatagegtt egettegegg 4500 ggtaaceaeg teteceeeae geactatgtg eetgagageg acgetgeage acgtgteaet 4560 cagateetet etagtettae cateaeteag etgetgaaga ggetteaeea gtggateaae 4620 gaggaetget ecaegeeatg eteeggeteg tggetaagag atgtttggga ttggataace 4680	accetectgt	LLAACATCCT	yyyyyyatgg	gradecacce	aactigotoo	coccagogor	
aaggtgettg tggatatttt ggeaggttat ggageagggg tggeaggege getegtggee 4320 tttaaggtea tgageggea gatgeeetee acegaggaee tggttaaeet acteeetget 4380 atceteteee etggegeeet agtegteggg gtegtgtgeg eagegataet gegteggeae 4440 gtgggeeeag gggagggge tgtgeagtgg atgaacegge tgatagegtt egettegegg 4500 ggtaaceaeg teteeeceae geactatgtg eetgagageg acgetgeage acgtgteaet 4560 cagateetet etagtettae eateacteag etgetgaaga ggetteaeea gtggateaee 4620 gaggaetget ecaegeeatg eteeggeteg tggetaagag atgtttggga ttggatatge 4680	gcttctgctt	tcgtaggcgc	cggcatcgct	ggagcggctg	ttggcagcat	aggccttggg	
tttaaggtca tgagcggcga gatgccctcc accgaggacc tggttaacct actccctgct 4380 atcctctccc ctggcgccct agtcgtcggg gtcgtgtgcg cagcgatact gcgtcgggac 4440 gtgggcccag gggaggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgcgg 4500 ggtaaccacg tctcccccac gcactatgtg cctgagagcg acgctgcagc acgtgtcact 4560 cagatcctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac 4620 gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatcatc 4680	aanntnotto	togatattt	aacsaattst	adsacsacaa	taggagggg	acteataace	4320
atcetetece etggegeeet agtegteggg gtegtgtgeg eagegataet gegteggeae 4440 gtgggeeeag gggagggge tgtgeagtgg atgaacegge tgatagegtt egettegegg 4500 ggtaaceaeg teteceeeae geactatgtg eetgagageg aegetgeage aegtgteaet 4560 eagateetet etagtettae eateaeteag etgetgaaga ggetteaeea gtggateaae 4620 gaggaetget eeaegeeatg eteeggeteg tggetaagag atgtttggga ttggatatge 4680	aagguguuu		Jacayartat	22~2~42333		J5505500	
atcetetece etggegeeet agtegteggg gtegtgtgeg eagegataet gegteggeae 4440 gtgggeeeag gggagggge tgtgeagtgg atgaacegge tgatagegtt egettegegg 4500 ggtaaceaeg tetececeae geactatgtg eetgagageg aegetgeage aegtgteaet 4560 eagateetet etagtettae eateaeteag etgetgaaga ggetteaeea gtggateaae 4620 gaggaetget eeaegeeatg eteeggeteg tggetaagag atgtttggga ttggatatge 4680	tttaaggtca	rgagcggcga	yatgccctcc	accgaggacc	iggitaacct	actecetget	
gtgggcccag gggaggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgcgg 4500 ggtaaccacg tctccccac gcactatgtg cctgagagcg acgctgcagc acgtgtcact 4560 cagatcctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac 4620 gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatatgc 4680	atcctctccc	ctggcgccct	agtcgtcggg	gtcgtataca	cagcgatact	gcgtcggcac	4440
ggtaaccacg tctccccac gcactatgtg cctgagagcg acgctgcagc acgtgtcact 4560 cagatcctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac 4620 gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatatgc 4680	ataaaaaaa	addeddaaaa	tataceatac	atraaccocc	tgatagggtt	cacttcacaa	4500
cagatcetet ctagtettae cateacteag etgetgaaga ggetteacea gtggateaac 4620 gaggaetget ecaegeeatg eteeggeteg tggetaagag atgtttggga ttggatatge 4680	gegggeedag	22242222	-aracaaraa	acgaaccygc	22424222	-33-39	
cagatectet etagtettae eateacteag etgetgaaga ggetteacea gtggateaac 4620 gaggaetget ecaegeeatg eteeggeteg tggetaagag atgtttggga ttggatatge 4680	ggtaaccacg	tctcccccac	gcactatgtg	cctgagagcg	acgctgcagc	acgigicact	
gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatatgc 4680	cagatectet	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gtggatcaac	4620
202200200 000020002002 022000200	daddadtadt	ccacacata	ctccccctcc	taactaacaa	atatttaaaa	ttggatatgg	4680
acggrgrrga crgarrrcaa gaccrggere cagrecaage recrgeegeg arrgeeggga 4740	gaggactgct	-t	Jedogge edg				
	acggtgttga	ctyatttcaa	yacctggctc	cagcccaagc	Lucigoogog	arryccygga	±/±0

	gtccccttct	tctcatgtca	acqtqqqtac	aagggagtct	qqcqqqqqqa	cggcatcatg	4800
	caaaccacct	gcccatgtgg	agcacagatc	accqqacatg	tgaaaaacgg	ttccatgagg	4860
	atcgtggggc	ctaggacctg	tagtaacacg	tggcatggaa	cattccccat	taacgcgtac	4920
	accacgggcc	cctgcacgcc	ctccccggcg	ccaaattatt	ctagggcgct	gtggcgggtg	4980
	gctgctgagg	agtacgtgga	ggttacgcgg	gtgggggatt	tccactacgt	gacgggcatg	5040
	accactgaca	acgtaaagtg	cccgtgtcag	gttccggccc	ccgaattctt	cacagaagtg	5100
	gatggggtgc	ggttgcacag	gtacgctcca	gcgtgcaaac	ccctcctacg	ggaggaggtc	5160
		tcgggctcaa					5220
	ccggacgtag	cagtgctcac	ttccatgctc	accgacccct	cccacattac	ggcggagacg	5280
	gctaagcgta	ggctggccag	gggatctccc	ccctccttgg	ccagctcatc	agctagccag	5340
		cttccttgaa					5400
		ccaacctcct					5460
		aggtagtaat					5520
		ccgttccggc					5580 5640
	tecatacggg	cacgcccgga	aggatataa	ttaacaata	agteetggaa	trantagae	5700
٠	catagedeece	cagtggtaca ggaagaggac	cgggtgtcca	tcacaatcta	ccatatette	taccttaaca	5760 5760
	gaggtggga	caaagacctt	caccaactcc	gaatcgt	ccgtgcccc	caacacaaca	5820
	accocctctc	ctgaccagcc	ctccaacaac	aacascacaa	gatccgacgt	tgagtcgtac	5880
	tectecatee	cccccttga	aaaaasacca	agagataca	atctcacca	caaatcttaa	5940
	tctaccqtaa	gcgaggaggc	tagtgaggac	atcatctact	gctcgatgtc	ctacacatqq	6000
	acagggggcc	tgatcacgcc	atgcactaca	qaqqaaacca	agctgcccat	caatgcactg	6060
		tgctccgtca					6120
		agaaggtcac					6180
	gtgctcaagg	agatgaaggc	gaaggcgtcc	acagttaagg	ctaaacttct	atccgtggag	6240
	gaagcctgta	agctgacgcc	cccacattcg	gccagatcta	aatttggcta	tggggcaaag	6300
	gacgtccgga	acctatccag	caaggccgtt	aaccacatcc	gctccgtgtg	gaaggacttg	6360
		ctgagacacc					6420
	gtccaaccag	agaaggggg	ccgcaagcca	gctcgcctta	tcgtattccc	agatttgggg	6480
	gttcgtgtgt	gcgagaaaat	ggccctttac	gatgtggtct	ccaccctccc	tcaggccgtg	6540
	atgggctctt	catacggatt	ccaatactct	cctggacagc	gggtcgagtt	cctggtgaat	6600 6660
	gcctggaaag	cgaagaaatg	ccctatgggc	anatanatat	acacccgctg	tazattagaa	6720
	acggteactg	agaatgacat gacaggccat	aggregate	agreage	tttacatcoo	agacteggee	6780
	actaattata	aagggcagaa	ctacaactet	caccaatacc	acacaacaa	tatactaaca	6840
	accaactaca	gtaataccct	cacatottac	ttgaaggccg	ctacaaccta	tcgagctgcg	6900
	aaactccaaa	actgcacgat	actcatatac	ggagaggagg	ttatcattat	ctgtgaaagc	6960
	acaaaaaccc	aagaggacga	gacgagecta	cgggccttca	cggaggctat	gactagatac	7020
	tetgecece	ctggggaccc	gcccaaacca	gaatacgact	tggagttgat	aacatcatgc	7080
	tcctccaatg	tgtcagtcgc	gcacgatgca	tctggcaaaa	gggtgtacta	tctcacccgt	7140
	gaccccacca	cccccttgc	gcgggctgcg	tgggagacag	ctagacacac	tccagtcaat	7200
	tcctggctag	gcaacatcat	catgtatgcg	cccaccttgt	gggcaaggat	gatcctgatg	, 7260
		tctccatcct					7320
		cctgttactc					7380
	catggcctta	gcgcattttc	actccatagt	tactctccag	gtgagatcaa	tagggtggct	7440
	tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	7500
		ggctactgtc					7560 7620
	aactgggcag	taaggaccaa	gereaaaere	accccaatcc	tatatgagag	agtatataat	7680
	cracecage	ggttcgttgc gctggttcat	atactacata	gygygagaca	ctatacacay	aggratetat	7740
	gecegaeeee	accgatgaac	gragaractaa	acactccacc	ccaataggggc	atcctatttt	7800
	tttccctttt	ttttttttt	ttttttt	tttttttt	tttttttt	ttctcctttt	7860
	tttttcctct	ttttttcctt	ttettteett	taataactcc	atcttagccc	tagtcacggc	7920
		aggtccgtga					7980
	agatcaagta	ctcctgcagg	cgcgccacta	gtgggaatac	gcggggtatg	ccgcgtttta	8040
		gacccaattc					8100
		agttaaattg					8160
		tcctcggcac					8220
	gtactgccgg	gcctcttgcg	ggatatcgtc	cattccgaca	gcatcgccag	tcactatggc	8280
	gtgctgctag	cgctatatgc	gttgatgcaa	tttctatgcg	cacccgttct	cggagcactg	8340
	tccgaccgct	ttggccgccg	cccagtcctg	ctcgcttcgc	tacttggagc	cactatcgac	8400
		tggcgaccac					8460
		gcgccacagg					8520
	gaagatcggg	ctcgccactt	cgggctcatg	agegettgtt	reggegtggg	tatggtggca	8580
		ccgggggact					8640
	geggtgetea	acggcctcaa	ectactactg	ggctgcttcc	caacgcagga	grogoataag	8700

PCT/US2003/039722 WO 2004/055216

agagagcatc	gaccgatgcċ	cttgagagcc	ttcaacccag	tcagctcctt	ccggtgggcg	8760
cadaacataa	ctatcatcac	cacacttata	actotettet	ttatcatgca	actcqtaqqa	8820
cagataccag	caacactcta	ggtcattttc	ggcgaggacc	gctttcgctg	gagcgcgacg	8880
atgatcggcc	tatcacttac	ggtattcgga	atcttgcacg	ccctcgctca	agccttcgtc	8940
actootccco	ccaccaaacq	tttcqqcqaq	aaqcaqqcca	ttatcgccgg	catggcggcc	9000
gacgcgctgg	actacatett	actagacatta	gcgacgcgag	gctggatggc	cttccccatt	9060
atgattcttc	tcacttccaa	caacatcaaa	atgcccqcqt	tgcaggccat	gctgtccagg	9120
caggtagatg	accaccatca	qqqacaqctt	caaggatcgc	tcgcggctct	taccagccta	9180
acttcgatca	ctggaccgct	gatcgtcacg	gcgatttatg	ccgcctcggc	gagcacatgg	9240
aacqqqttqq	catggattgt	aggegeegee	ctataccttg	tatgaataca	cgcgttgcgt	9300
cacaatacat	adadccaddc	cacctcgacc	tgaatggaag	ccggcggcac	ctcgctaacg	9360
gattcaccac	tccaaqaatt	ggagccaatc	aattcttgcg	gagaactgtg	aatgcgcaaa	9420
ccaacccttq	qcaqaacata	tccatcgcgt	ccgccatctc	cagcagccgc	acgcggcgca	9480
tctcqqqcaq	cattagatec	tggccacggg	tgcgcatgat	cgtgctcctg	tcgttgagga	9540
cccaactaaa	ctaacaaaat	tgccttactg	qttaqcaqaa	tgaatcaccg	atacgcgagc	9600
gaacgtgaag	caactactac	tgcaaaacgt	ctqcqacctq	agcaacaaca	tgaatggtct	9660
tcaatttcca	totttcotaa	agtctggaaa	cacagaagtc	agcgccctgc	accattatgt	9720
tecagateta	caticocagga	tactactaac	taccctqtqq	aacacctaca	tctqtattaa	9780
casaacacta	gcattgaccc	tgagtgattt	ttctctqqtc	ccgccgcatc	cataccgcca	9840
attatttacc	ctcacaacqt	tccactaacc	gggcatgttc	atcatcagta	acccqtatcq	9900
tgagcatcct	ctctcatttc	atcogtatca	ttacccccat	gaacagaaat	tcccccttac	9960
acqqaqqcat	caaqtqacca	aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
aagggagaga	ttaacgcttc	togagaaact	caacgagctg	gacgcggatg	aacaggcaga	10080
catctgtgaa	tcacttcaca	accacqctqa	tgagctttac	cgcagctgcc	tcgcgcgttt	10140
caataataac	gatgaaaacc	tctgacacat	gcagctcccg	gagacggtca	cagcttgtct	10200
gtaagcggat	accadaaaca	gacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	10260
t.cggggggga	gccatgaccc	agtcacgtag	cqataqcqqa	gtgtatactg	gcttaactat	10320
gcggcatcag	agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tacataaaaa	gaaaataccg	catcaggcgc	tcttccqctt	cctcgctcac	tgactcgctg	10440
cactcaatca	ttcaactaca	acaaacaata	tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat	cagggataa	cacadaaaa	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaggccgc	attactaaca	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgacgctc	aaqtcaqaqq	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	cccctggaag	ctccctcata	cactatacta	ttccgaccct	gccgcttacc	10740
ggatacctgt	ccacctttct	cccttcggga	agcataacac	tttctcatag	ctcacgctgt	10800
aggtatctca	attcaatata	aatcattcac	tccaaqctqq	gctgtgtgca	cgaacccccc	10860
atteagecea	accactacac	cttatccqqt	aactatcqtc	ttgagtccaa	cccggtaaga	10920
cacgacttat	caccactage	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	10980
aacaatacta	cagagttett	gaagtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttggtatct	acactctact	gaagccagtt	accttcqqaa	aaagagttgg	tagctcttga	11100
tccggcaaac	aaaccaccgc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
cocagaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	11220
togaacgaaa	actcacqtta	aggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
tagatccttt	tctagataat	acgactcact	ata	· · · · · ·		11313
						,
	0> 9					
	1> 11313					
	2> DNA	nl Comicnes				
<21.	3> Artificia	ar peduence				

```
<220>
<223> Plasmid
```

<400> 9

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120 180 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc 240 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300 360 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 420 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 480 cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540 600

acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca

cgacqqqcqt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660 tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720 aagtatccat catggetgat geaatgegge ggetgeatae gettgateeg getaeetgee 780 cattcgacca ccaagcgaaa catcgcatcg agcgagcacg tactcggatg gaagccggtc 840 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900 ccaggeteaa ggegegeatg eccgaeggeg aggatetegt egtgaeceat ggegatgeet 960 gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1020 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080 ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140 agogoatogo ottotatogo ottottgaog agitottotg agittaaaca gaccacaacg 1200 gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260 1320 cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1380 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct aggggtettt cecetetege caaaggaatg caaggtetgt tgaatgtegt gaaggaagea 1440 gttcctctgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1500 aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560 1620 gcaaaggcgg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggt accccattgt 1680 1740 atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gaggttaaaa 1800 aacgtctagg cccccgaac cacggggacg tggttttcct ttgaaaaaca cgataatacc 1860 atggegeeta ttaeggeeta eteceaacag aegegaggee taettggetg cateateaet 1920 agcctcacag gccgggacag gaaccaggtc gagggggagg tccaagtggt ctccaccgca acacaatett teetggegae etgegteaat ggegtgtgtt ggaetgteta teatggtgee 1980 ggctcaaaga cccttgccgg cccaaagggc ccaatcaccc aaatgtacac caatgtggac 2040 2100 caggaceteg teggetggea agegeeece ggggegegtt cettgaeace atgeacetge 2160 ggcagctcgg acctttactt ggtcacgagg catgccgatg tcattccggt gcgccggcgg 2220 qqcqacaqca qqqqqaqcct actctccccc aggcccgtct cctacttgaa gggctcttcg 2280 ggcggtccac tgctctgccc ctcggggcac gctgtgggca tctttcgggc tgccgtgtgc acccgagggg ttgcgaaggc ggtggacttt gtacccgtcg agtctatggg aaccactatg 2340 cggtccccgg tcttcacgga caactcgtcc cctccggccg taccgcagac attccaggtg 2400 gcccatctac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 2460 2520 gcccaagggt ataaggtgct tgtcctgaac cegtccgtcg ccgccaccct aggtttcggg 2580 gcgtatatgt ctaaggcaca tggtatcgac cctaacatca gaaccggggt aaggaccatc 2640 accaegggtg ecceeateae gtactecaee tatggeaagt ttettgeega eggtggttge tctgggggcg cctatgacat cataatatgt gatgagtgcc actcaactga ctcgaccact 2700 atcctgggca tcggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtcgtg 2760 ctcgccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggaggtg 2820 2880 gctctgtcca gcactggaga aatccccttt tatggcaaag ccatccccat cgagaccatc aaggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccgcg 2940 3000 aagctgtccg gcctcggact caatgctgta gcatattacc ggggccttga tgtatccgtc 3060 ataccaacta geggagacgt cattgtegta gcaacggacg ctctaatgac gggctttacc 3120 ggcgatttcg actcagtgat cgactgcaat acatgtgtca cccagacagt cgacttcagc ctggacccga ccttcaccat tgagacgacg accgtgccac aagacgcggt gtcacgctcg 3180 cagcggcgag gcaggactgg taggggcagg atgggcattt acaggtttgt gactccagga 3240 gaacggccct cgggcatgtt cgattcctcg gttctgtgcg agtgctatga cgcgggctgt 3300 3360 gcttgqtacq agctcacgcc cgccgagacc tcagttaggt tgcgggctta cctaaacaca 3420 ccagggttgc ccgtctgcca ggaccatctg gagttctggg agagcgtctt tacaggcctc acccacatag acgcccattt cttgtcccag actaagcagg caggagacaa cttcccctac 3480 ctggtagcat accaggctac ggtgtgegce agggeteagg etecacetec ategtgggae 3540 caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gcccctgctg 3600 tataggctgg gagccgttca aaacgaggtt actaccacac accccataac caaatacatc 3660 3720 atggcatgca tgtcggctga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 3780 gtcctagcag ctctggccgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg 3840 atcatcttgt ccggaaagcc ggccatcatt cccgacaggg aagtccttta ccgggagttc 3900 gatgagatgg aagagtgcgc ctcacacetc cettacateg aacagggaat geagetegee gaacaattca aacagaaggc aatcgggttg ctgcaaacag ccaccaagca agcggaggct 3960 gctgctcccg tggtggaatc caagtggcgg accctcgaag ccttctgggc gaagcatatg 4020 tggaatttca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080 4140 gcgatagcat cactgatggc attcacagcc tctatcacca gcccgctcac cacccaacat 4200 accetectgt ttaacateet ggggggatgg gtggeegeec aaettgetee teecageget gcttctgctt tcgtaggcgc cggcatcgct ggagcggctg ttggcagcat aggccttggg 4260 4320 aaggtgettg tggatatttt ggeaggttat ggageagggg tggeaggege getegtggee tttaaggtca tgagcggcga gatgccctcc accgaggacc tggttaacct actccctgct 4380 atcetetece etggegeet agtegteggg gtegtgtgeg eagegataet gegteggeae 4440 gtgggcccag gggaggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgcgg 4500 4560 ggtaaccacg tetececeae geactatgtg cetgagageg acgetgeage acgtgteact

cacatectet	ctagtettae	catcactcac	ctactaaaaa	ggcttcacca	qtqqatcaac	4620
cagacecece	Coagoood		50900900		+++	
gaggactgct	ccacgccatg	ctccggctcg	tggctaagag	atgtttggga	Liggalatge	4680
5 55 5 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	atastttass	gacctgggtc	cactccaacc	tcctgccgcg	attaccagaa	4740
acggiginga	Cigacicaa	gacceggeee	cageccaage	2000300303	33-3-	
atccccttct	tctcatgtca	acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	4800
	aggatatag	addadata	acconsta	tgaaaaacgg	ttccatgagg	4860
Caaaccacci	geceatgegg	agcacagacc	accegacace	2944444099	cccacgagg	
atcotogooc	ctaggacctg	tagtaacacg	tqqcatqqaa	cattecccat	taacgcgtac	4920
~~5~5555				atzaaaaaat	ataaaaata	4980
accacgggcc	cctgcacycc	cccccggcg	CCaaactatt	ctagggcgct	9-99-999-9	
actactasaa	agtacgtgga	aattacacaa	ataaaaaatt	tccactacgt	gacgggcatg	5040
5005005055	ugum-g-gg-	33			2 222 2	5100
accactgaca	acgtaaagtg	cccgtgtcag	grreeggeee	ccgaattctt	cacagaagig	
astagastaa	aattacacaa	atacacticca	acatacaaac	ccctcctacg	agaagaatc	5160
gacggggcgc	ggccgcacag	500050000	303030000			
acattcctqq	tegggeteaa	tcaatacctg	gttgggtcac	agctcccatg	cgagcccgaa	5220
aaaaaaataa	anataatana	ttacatacta	accoacccct	cccacattac	aacaaaaaca	5280
deggaegtag	Caguguucac	cccacgccc	accgaccccc		3303343403	
actaaacata	aactaaccaa	gggatctccc	ccctccttgg	ccagctcatc	agctatccag	5340
55-5		222	actacacta	atractoro	anacactasa	5400
ctgtctgcgc	cttccttgaa	ggcaacatgc	actaccegic	atgactcccc	ggacgccgac	
ctcatcgagg	ccaacctcct	ataacaacaa	gagatggggg	ggaacatcac	ccgcgtggag	5460
			++ = = = = = = = = = = = = = = = = = =	+ ada adddda	aaaaaataaa	5520
tcagaaaata	aggragraar	cciggagici	LLegageege	tccaagcgga	ggaggargag	
adddaadtat	ccattccaac	ggagatcctg	caaaaatcca	ggaaattccc	tcgagcgatg	5580
agggaageae	0090000990	33-3			~~~~~~~~~~	5640
cccatatggg	cacgcccgga	ttacaaccct	ccactgttag	agtcctggaa	ggacccggac	
tacatacata	cartrontaca	caaatataca	ttaccaccta	ccaaggcccc	tccgatacca	5700
Lacgicelle	cageggeaca	099909000			<u> </u>	5760
cctccacgga	ggaagaggac	ggttgtcctg	tcagaatcta	ccgtgtcttc	rgeerrggeg	
asaataaaas	da aadacctt	caacaactcc	gaatcgtcgg	ccgtcgacag	caacacaaca	5820
gagetegeta	Caaagacccc	caacca	9445095099	005005005		E000
acqqcctctc	ctgaccagcc	ctccgacgac	ggcgacgcgg	gatccgacgt	tgagtegtae	5880
++	aaaaaattaa	adddd a dd dd	agagetaca	atctcagcga	cagatettaa	5940
teeteeatge	ecceccinga	gggggagccg	ggggaccccg	acccoagoga	0999000033	
tctaccqtaa	acaaaaaaac	tagtgaggac	gtcgtctgct	gctcgatgtc	ctacacatgg	6000
000000	5-5-55-55-		222222	agatagagat	gaatggagtg	6060
acaggcgccc	tgateaegee	argegergeg	gaggaaacca	agctgcccat	caacgcaccg	
aggaagtett	tactccatca	ccacaactto	atctatacta	caacatctcg	cagcgcaagc	6120
ugouuocco.	230000			+~~~~~~~	at a addada a	6180
ctgcggcaga	agaaggtcac	etttgacaga	clycaggicc	tggacgacca	Ctaccyggac	
atactcaaaa	adatdaaddc	gaaggggtcc	acaqttaaqq	ctaaacttct	atccqtqqaq	6240
90900000	4940944990	54455-5		++	+aaaaaaaaaa	6300
gaagcctgta	agctgacgcc	cccacattcg	gecagateta	aatttggcta	Lggggcaaag	
ascat acads	acctatccad	caaggccgtt	aaccacatcc	gctccgtgtg	gaaggacttg	6360
gacgcccgga	ucceucoug	0445500500			+++-	6420
ctggaagaca	ctgagacacc	aattgacacc	accatcatgg	caaaaaatga	ggttttetge	
ataasaasa	adaaddddd	ccacaaacca	actcacctta	tcgtattccc	agatttgggg	6480
gcccaaccag	agaaggggg	cegeaageea	300030000			6540
atteatatat	gcgagaaaat	ggccctttac	gatgtggtct	ccaccctccc	teaggeegig	
atagactatt	catacqqatt	ccaatactct	cctggacagc	gggtcgagtt	cctqqtqaat	6600
atgggctctt	cacacggacc	CCAACACCCC	0009940490	3333		
qcctqqaaaq	cgaagaaatg	ccctatgggc	ttcgcatatg	acacccgctg	ttttgaetea	6660
pagatanata	agaatgagat	ccatattasa	cactcaatct	accaatgttg	tgacttggcc	6720
acggicacig	agaatgacat	ccgcgccgug	gagccaaccc	accadegacg		
cccgaagcca	qacaqqccat	aaggtcgctc	acagagcggc	tttacatcgg	gggcccccrg	6780
	2 22	atagaaatat	caccaatacc	gcgcgagcgg	tatactaaca	6840
actaattcta	aagggcagaa	Cigiggciai	cgccggcgcc	acacadasa		
accadctdcd	gtaataccct	cacatqttac	ttgaaggccg	ctgcggcctg	tcgagctgcg	6900
		~ataatataa	aasasaasaa	ttataattat	ctatassac	6960
aageteeagg	actycacyat	getegtatge	ggagacgacc	ttgtcgttat	cegegaaage	
acaaaaaccc	aaqaqqacqa	ggcgagccta	cgggccttca	cggaggctat	gactagatac	7020
t -t	ataaaaaaa	adddaaadda	gastacgact	tggagttgat	aacatcatoc	7080
Lordoccoc	ecggggaeee	gcccaaacca	gaacacgacc	caaraccarc	445454455	
tcctccaatq	tatcaatcac	gcacgatgca	tctggcaaaa	gggtgtacta	tctcacccgt	7140
		accedetace	taggagagag	ctenerecer	tccadtcaat	7200
gaccccacca	ececectige	gegggergeg	Lygyagacag	ctagacacac	ccagcaac	
tectooctag	gcaacatcat	catqtatqcq	cccaccttgt	gggcaaggat	gatcctgatg	7260
		tatacatana	annanatta	aaaaaaaaat	agattataag	7320
acteatttet	tetecatect	Lecayereay	gaacaacttg	aaaaagccct	agacogcoug	• -
atctacgggg	cctattactc	cattgagcca	cttgacctac	ctcagatcat	tcaacgactc	7380
		nataantnat	tagtatagaa	atasastass	tagggtgggt	7440
catggcctta	gegeattite	accedatage	Lacticicag	gtgagatcaa	cagggegge	
tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	7500
	~~~+~~+~+	ganaaaaaa	aggagtagg	cttgtggcaa	atacetette	7560
greegegera	ggeraergre	ccaggggggg	agggetgeta	cccgcggcaa	geacutous	
aactgggcag	taaqqaccaa	gctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	7620
	~~++~~++~~	taattaaaaa	aaaaaaaaaa	tatatcacag	cetateteat	7680
ttatccagct	ggtttgttgt	Lygicadage	gggggagaca	Lacaccacag	cccgccccgc	
acccaacccc	actaattcat	ataataccta	ctcctacttt	ctgtaggggt	aggcatctat	7740
	555	2222	agagtagaaa	dasatsaaaa	atactattt	7800
ctactcccca	accgatgaac	ggggacetaa	acactccagg	ccaataggcc	accorgance	
tttccctttt	ttttttttt	ttttttttt	ttttttttt	ttttttttt	ttctcctttt	7860
			+~~+~~	-+-+	tantnannn	7920
tttttcctct	LLLLLCCLL	LECEUTECCE	ragragatec	atcttagccc	cayccacyyc	- "
tageteteaa	addtccatas	accacttasc	tacadadadt	gctgatactg	gcctctctqc	7980
Laguegua	-55555554			~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	aaaaa++++	8040
agatcaagta	cccccgcagg	cgcgccacta	grgggaarac	gcggggtatg	cogoguetta	
acatattaac	gacccaattc	tcatotttca	cagettatea	tcgataagct	ttaatqcqqt	8100
Juliu Lugue					+	8160
agtttatcac	agrraaattg	ctaacgcagt	caggcaccgt	gtatgaaatc	Laacaacycg	
atastaatas	taataaaaaa	catcacacta	gatgetgtag	gcataggctt	ggttatgccg	8220
Juliungua	2000099040		J J J J			
gtactgccgq	gcctcttgcq	ggatatcgtc	cattccgaca	gcatcgccag	ceactatgge	8280
atantantan	cactatataa	attastacss	tttatataaa	daddddttdt	cggagcactg	8340
guguuguag	Jecalaly	July Lugena		100000000000000000000000000000000000000		
tccgaccgct	ttggccgccq	cccagtcctg	ctcgcttcgc	tacttggagc	cactatcgac	8400
+		, -	<u> </u>			8460
	taaaaaaaaa	acccut coto	LUCSTCCTCT	acaccacaco	Cattuluce	0400
tacgegatea	tggcgaccac	accegteetg	tggatcctct	acgccggacg	categuggee	
ggcatcaccq	tggcgaccac gcgccacagg	tgcggttgct	ggcgcctata	tcgccgacat	caccgatggg	8520

gaagatcggg	ctcqccactt	cgggctcatg	agcgcttgtt	tcggcgtggg	tatggtggca	8580
aaccccataa	ccagagaact	gttgggcgcc	atctccttqc.	atgcaccatt	ccttgcggcg	8640
acaatactca	acooctcaa	cctactactg	gactacttcc	taatgcagga	gtcgcataag	8700
ggagaggata	gaccgatgcc	cttgagagcc	ttcaacccaq	tcagctcctt	ccggtgggcg	8760
cadadcatas	ctatcatcac	cgcacttatg	actatcttct	ttatcatgca	actcgtagga	8820
canataccan	caacactcta	ggtcattttc	gacgaggacc	actttcacta	gagegegaeg	8880
atratronco	tatcacttac	ggtattcgga	atcttgcacg	ccctcqctca	agccttcatc	8940
actootoco	ccaccaaaco	tttcggcgag	aaggaggga	ttatcgccgg	catageagee	9000
accedence	actacatett	gctggcgttc	acascacasa	actagatage	cttccccatt	9060
atrattotto	tractteraa	cggcatcggg	atacccacat	tacagaccat	actatccaaa	9120
caratagata	accaccatca	gggacagctt	caaggatcgc	tcacaactct	taccagecta	9180
acttccatca	ctgaecact	gatcgtcacg	gcgatttatg	ccacctcaac	gagcacatgg	9240
accedated	catocattot	aggcgccgcc	ctataccttq	tetacetece	cacattacat	9300
aacgggccgg	cacadacada	cacctcgacc	taaataaaaa	ccaacaacsc	ctcactaaca	9360
cattangan	taaaaaatt	ggagccaatc	aattettee	gagaactgtg	aatgcgcaaa	9420
gatteaceac	cccaagaatt	tccatcgcgt	ccccatctc	dadaaaaaaaa	acacaacaca	9480
tehagggang	gcagaacaca	tggccacggg	tagacatast	catactacta	tcattaaaaa	9540
LCLCgggCag	cyclygyldd	tgccttactg	attaggac	tastasca	atacccaacc	9600
cccggctagg	ccggcggggc	tgcaaaacgt	gttagtagaa	accaacaaca	traatrotet	9660
gaacgtgaag	egaetgetge	agtctggaaa	agaggaectg	agcaacaaca	accattatot	9720
teggttteeg	tgtttcgtaa	agrurggaaa	taggggaagte	agegeeeege	tctatattaa	9780
teeggatetg	cacegeagga	tgctgctggc	taccetgegg	aacacccaca	cataccacca	9840
cgaagcgctg	gcattgaccc	tgagtgattt		atantanata	agggtateg	9900
gttgtttacc	ctcacaacgt	tccagtaacc	gggcatgttc	accaccagca	tagagettag	9960
tgagcatcct	ctctcgtttc	atcggtatca	ttacccccat	yaacayaaac	agtttataag	10020
acggaggcat	caagtgacca	aacaggaaaa	aacegeeett	aacatggccc	getttateag	10020
aagccagaca	ttaacgcttc	tggagaaact	caacgagerg	gaegeggaeg	aacaggcaga	10140
catctgtgaa	togetteacg	accacgctga	tgagctttac	cgcagctgcc	regegegeee	10240
cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	cagettgtet	
gtaagcggat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	rrggcgggrg	10260
tcggggcgca	gccatgaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	10320
gcggcatcag	agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	10380
tgcgtaagga	gaaaataccg	catcaggcgc	tcttccgctt	cctcgctcac	tgactcgctg	10440
cgctcggtcg	ttcggctgcg	gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	10500
tccacagaat	caggggataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	10740
ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	10800
aggtatctca	gttcggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	10860
gttcagcccg	accgctgcgc	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	10920
cacgacttat	cqccactqqc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	10980
ggcggtgcta	cagagttctt	gaagtggtgg	cctaactacg	gctacactag	aaggacagta	11040
tttqqtatct	gegetetget	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	11100
tccqqcaaac	aaaccaccgc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
cacagaaaaa	aaqqatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	11220
tagaacaaaa	actcacqtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
tagateettt	tctagataat	acgactcact	ata			11313
٥	-	-				

```
<210> 10
<211> 11313
<212> DNA
<213> Artificial Sequence
<220>
<223> Plasmid
```

### <400> 10

gccagcccc gattggggc gacactccac catagatcac tecectgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
cccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgcccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaaggg cgcgcatga ttgaacaaga tggattgcac gcaggttctc 420

					~+~~~~+	400
cggccgcttg	ggrggagagg	CLattegget	acgaccgggc	acaacagaca	accygerger	480
ctgatgccgc	catattccaa	ctqtcaqcqc	aggggggcc	ggttcttttt	gtcaagaccg	540
						600
				gcggctatcg		
cgacgggcgt	teettgegea	gctgtgctcg	acqttqtcac	tgaagcggga	agggactqqc	660
+00+0++	aassataaaa	adadadata	taatatasta	tcaccttgct	cctacaaaa	720
rgccarrggg	cgaagegeeg	gggcaggacc	cccegccacc	coaccegee	ccegacgaga	
aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	gcttgatccg	gctacctgcc	780
asttaasaas	ccaaccaaa	catcacatca	adcdaddadd	tactcggatg	gaagecoote	840
Callegadea	ccaagegaaa	caccycaccy	agegageacg	cacceggaeg	gaageeggee	
ttqtcqatca	ggatgatctg	gacgaagagc	atcaggggct	cgcgccagcc	gaactgttcg	<sup>'</sup> 900
adadatasa	gagagagata	accas accaca	aggatetegt	cgtgacccat	aacastacct	960
Ccaggcccaa	ggcgcgcarg	cccgacggcg	aggatetege	cgcgacccac	ggcgacgccc	
gcttgccgaa	tatcatggtg	gaaaatggcc	gcttttctgg	attcatcgac	tgtggccggc	1020
tagatataga	agagggtat	caccacatac	cattaactac	ccgtgatatt	actassasa	1080
ttggcggcga	atgggctgac	cgcttcctcg	tgctttacgg	tatcgccgct	cccgattcgc	1140
adddatada	attatataca	cttcttcacc	agttetteta	agtttaaaca	garcacaacg	1200
agegeatege	CCCCCaccgc	cccccgacg	ageceeeg	ageceaaaca	gaccacaacg	
gtttccctct	agcgggatca	attccgcccc	tatacataca	cccccctaa	cgttactggc	1260
cassaccact	tagaataaga	ccaatataca	tttqtctata	tgttattttc	caccatatto	1320
cgaageegee	caaaaaaa	0099090909		b-65-65-6-		
ccgtcttttg	gcaatgtgag	ggcccggaaa	cctggccctg	tcttcttgac	gageatteet	1380
aggggtcttt	cccatataga	caaaggaatg	caaggtetgt	tgaatgtcgt	gaaggaagga	1440
499990000					3333	
				cgaccctttg		1500
aaccccccac	ctaacaacaa	atacetetae	ggccaaaagc	cacgtgtata	agatacacct	1560
						1620
				tagttgtgga		
taactctcct	caagcgtatt	caacaaqqqq	ctgaaggatg	cccagaaggt	accccattgt	1680
						1740
atgggatetg	arcrggggce	teggtgeaea	egetttaeat	gtgtttagtc	gaggitaaaa	
aacqtctagq	ccccccaac	cacaaaaaaa	taattttcct	ttgaaaaaca	cgataatacc	1800
	++	ataganaa		trattaaata	antantanat	1860
atggegeeta	ttacggccta	CLCCCaacag	acgegaggee	tacttggctg	Cattattatt	
agceteacag	accaaaacaa	gaaccaggtc	aaaaaaaaa	tccaagtggt	ctccaccqca	1920
	++	ataaataaat	accetatat	gazatatata	tastaataaa	1980
acacaatett	ceetggegae	Cigogicaac	ggcgcgcgcc	ggactgtcta	ccacggcgcc	
gactcaaaga	cccttaccaa	cccaaaqqqc	ccaatcaccc	aaatgtacac	caatgtggac	2040
22	tarastaras	200000000	aaaaaaaatt	aattaaaaa	atacaactac	2100
caggaceteg	reggerggea	agegeeeeee	ggggcgcgcc	ccttgacacc	acgeaceege	
ggcagctcgg	acctttactt	ggtcacgagg	catgccgatg	tcattccggt	gcgccggcgg	2160
addasasas	aaaaaaaaat	agtatacaca	aggecentet	cctacttgaa	agactettea	2220
ggcgacagca	gggggagccc	accecece	aggeeegeee	cccaccegaa	99900000	
ggcggtccac	tgctctgccc	ctcggggcac	gctgtgggca	tctttcgggc	tgccgtgtgc	2280
3000030000	ttacasaaa	aataaacttt	atacceatea	agtctatggg	aaccactato	2340
accegaggg	ccacagaaaac	aacaaacccc	geacecgeog		aaccaccacg	
cggtccccgg	tcttcacgga	caactcgtcc	cctccggccg	taccgcagac	attccaggtg	2400
				aggtgccggc		2460
gcccaagggt	ataaggtgct	tgtcctgaac	ccgtccgtcg	ccgccaccct	aggetteggg	2520
acatatatat	chaaggcaca	togtatogac	cctaacatca	gaaccggggt	aaggaccatc	2580
accacgggtg	ccccatcac	gtactccacc	tatggcaagt	ttcttgccga	eggeggeege	2640
tetagagaga	cctatoacat	cataatatot	gatgagtgcc	actcaactga	ctcgaccact	2700
				ctggagcgcg		2760
ct.cgccaccg	ctacccctcc	aaaatcaatc	accataccac	atccaaacat	caaqqaqqtq	2820
						2880
getetgteea	gcactggaga	aaceeectt	Latyycaaay	ccatccccat	cgagaccatc	
aaqqqqqqa	ggcacctcat	tttctqccat	tccaaqaaqa	aatgtgatga	gctcgccgcg	2940
222222	aaataaaaat	anataatata	acatattaca	aaaaaattaa	tatatacata	3000
aagergreeg	geeleggael	caatgutgta	gcatattact	ggggccttga	LgLaccegee	
ataccaacta	gcggagacgt	cattqtcqta	gcaacggacg	ctctaatgac	gggctttacc	3060
~~~~	5-55 5 5	agaataaaat	agatatataa	aaaaaaaa	agaattaaga	3120
ggcgattteg	acteagegat	cgactgcaat	acacycycca	cccagacagt	cgaccccage	
ctggacccga	ccttcaccat	tgagacgacg	accgtgccac	aagacgcggt	gtcacgctcg	3180
cadcadcaaa	acadactaa	tagggggggg	atogocattt	acaggtttgt	gactccagga	3240
cagoggogag	9049940699	0099990099			5	
gaacggccct	cgggcatgtt	cgattcctcg	gttctgtgcg	agtgctatga	cgcgggctgt	3300
acttaatea	ageteacgee	caccaagacc	t.cagt.t.aggt	tgcgggctta	cctaaacaca	3360
3000330003	250000500	-55		44444	h	3420
ccagggttgc	cegtetgeca	ggaccatctg	gagttetggg	agagcgtctt	caeaggeere	
acccacatag	acocccattt	cttatcccaa	actaaqcaqq	caggagacaa	cttcccctac	3480
-44				at accept ac	atactacaaa	3540
				ctccacctcc		
caaatqtqqa	agtgtctcat	acqqctaaaq	cctacqctqc	acgggccaac	gcccctgctg	3600
						3660
				accccataac		
atggcatgca	tgtcagctga	cctggaggtc	gtcacqaqca	cctgggtgct	ggtaggcgga	3720
				gcgtggtcat		3780
atcatcttgt	ccggaaagcc	ggccatcatt	cccgacaggg	aagtctttta	ccgggagttc	3840
				aacggggaat		3900
gaacatttca	aacagaaggc	aatcgggttg	ctgcaaacag	ccaccaagca	agcggaggct	3960
				ccttctgggc		4020
rggaatttca	rcagcgggat	acaatattta	gcaggcttgt	ccactctgcc	rggcaacccc	4080
				gcccgctcac		4140
accetectgt	ttaacatcct	gggggatgg	yrggccgccc	aacttgctcc	Leccageget	4200
				ttggcagcat		4260
						4320
aaygtgcttg	uggatatttt			tggcaggcgc		
						4222
	tgagcggcga	gatgccctcc	accqaqqacc	tggttaacct	actccctqct	4380

atcctctccc	ctggcgccct	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4440
apacadaaa	gggagggggc	tataceataa	atraaccooc	tgatagggtt	cacttcacaa	4500
gragacceag	999999999	Lg Lg Lag Lgg	acgaaccggo.		0300003033	4560
ggtaaccacg	tctccccac	gcactatgtg	cctgagagcg	acgergeage	acgigicaci	
cagatcctct	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gtggatcaac	4620
garactact	ccacgccatg	checaachea	tooctaagag	atatttaaaa	ttagatatac	4680
gaggaccgcc	ccacgccacg	acatacata	andtachard	tactacaca	attaccaaaa	4740
acggtgttga	ctgatttcaa	gaddlygdld	cageccaage	ccccgccgcg	accyccyyga	
gtccccttct	tctcatgtca	acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	4800
caaaccacct	gcccatgtgg	agcacagatc	accggacatg	tgaaaaacgg	ttccatgagg	4860
staataaaa	ctaggacctg	tantaararn	tagcatagaa	cattccccat	taacqcqtac	4920
accgryyggc	ccaggacceg	Lagraacacg		ataccccat	ataaaaaata	4980
accacgggcc	cctgcacgcc	ereceeggeg	CCaaactact	ctagggtget	grggrgggrg	
gctgctgagg	agtacgtgga	ggttacgcgg	gtgggggatt	tccactacgt	gacgggcatg	5040
accactgaca	acgtaaagtg	cccatatcaa	gttccggccc	ccgaattctt	cacagaagtg	5100
astacasta	ggttgcacag	atacactaca	acatacasac	cochactaca	agaggaggt.c	5160
gatggggtgt	ggccgcacag	beek	9090904440	aggragata	22-22-22-	5220
acattcctgg	tcgggctcaa	teaatacetg	grigggread	aguitudaty	cyaycccyaa	
ccggacgtag	cagtgctcac	ttccatgctc	accgacccct	cccacattac	ggcggagacg	5280
actaaacata	ggctggccag	gggatctccc	ccctccttgg	ccagctcatc	agctagccag	5340
atatataaa	cttccttgaa	agcaacatac	actacccqtc	atoactcccc	ggacgctgac	5400
cigicigege		9904404090	anantagga	adagataaa	ccacataasa	5460
ctcatcgagg	ccaacctcct	grggeggeag	gagargggcg	ggaacaccac	ccgcgcggag	
tcagaaaata	aggtagtaat	tttggactct	ttcgagccgc	tccaagcgga	ggaggatgag	5520
agggaagtat	ccgttccggc	ggagatcctg	cggaggtcca	ggaaattccc	tcgagcgatg	5580
aggetetaga	cacgcccgga	ttacaaccct	ccactgttag	agtectggaa	ggacccggac	5640
cccatatggg	cacgeeegga		ttaaaaaata	adaadaaaa	toccatacca	5700
tacgtccctc	cagtggtaca	egggtgteea	Ligitogetty	ccaaggcccc	t	
cctccacgga	ggaagaggac	ggttgtcctg	tcagaatcta	ccgtgtcttc	tgccttggcg	5760
gagetegeea	caaagacctt	caacaactcc	gaatcgtcgg	ccgtcgacag	cggcacggca	5820
agggggggg	ctgaccagcc	ctcccaccac	aucascacaa	gatccgacgt	tgagtcgtac	5880
acggeerere	Cigactagee	Ccccgacgac	ggcgacgcgg	atatasaaa	acceptattee	5940
tectecatge	cccccttga	gggggagccg	ggggateeeg	accidagega	cyggictigg	
tctaccgtaa	gcgaggaggc	tagtgaggac	gtcgtctgct	gctcgatgtc	ctacacatgg	6000
acadddddddd	tgatcacgcc	atacactaca	gaggaaacca	agetgeecat	caatgcactg	6060
	tgctccgtca	acacaactta	atctatacta	caacatctco	cadcacaadc	6120
ageaactett	Lycholycoa	Ccacaacceg	-b		at a daddaaa	6180
ctgcggcaga	agaaggtcac	ctttgacaga	crgcaggree	Lygacyacca	CLaccyggac	
gtgctcaagg	agatgaaggc	gaaggcgtcc	acagttaagg	ctaaacttct	atccgtggag	6240
gaageetgta	agctgacgcc	cccacattcq	gccagatcta	aatttqqcta	tggggcaaag	6300
Saaseesse	acctatccag	caacccatt	aaccacatcc	acticiatata	gaaggacttg	6360
gacgcccgga	acctatccag	caaggeegee		3000030303	gattttata	6420
ctggaagaca	ctgagacacc	aattgacacc	accarcargy	Cadadaacya	ggttttttgt	
gtccaaccag	agaagggggg	ccgcaagcca	gctcgcctta	tcgtattccc	agatttgggg	6480
atteatatat	gcgagaaaat	aaccctttac	gatgtggtct	ccaccctccc	tcaggccgtg	6540
5tacgegege	catacggatt	ccaatactct	cctagacaga	gggtcgagtt	cctaataaat	6600
acgggererr	Catacygatt		therestate	999009090	ttttaatta	6660
gcctggaaag	cgaagaaatg	ccctatgggc	tregearary	acacccgctg	LLLLyacica	
acggtcactg	agaatgacat	ccgtgttgag	gagtcaatct	accaatgttg	tgacttggcc	6720
cccaaaacca	gacaggccat	aaggtcgctc	acagagegge	tttacatcgg	gggccccctg	6780
	aagggcagaa	atacaactat	caccaatacc	acacasacaa	tatactaaca	6840
actaatteta	aagggcagaa	cegeggeeac	-b	atacacata	tacaactaca	6900
accagctgcg	gtaataccct	cacatgttac	ttgaaggeeg	ergeggeerg	Legagergeg	
aagctccagg	actgcacgat	gctcgtatgc	ggagacgacc	ttgtcgttat	ctgtgaaagc	6960
acaaaaaccc	aagaggacga	ggcgagccta	cgggccttca	cggaggctat	gactagatac	7020
50555550000	ctggggaccc	acccasacca	gaatacgact	tagaattaat	aacatcatqc	7080
Letycoccc	ceggggacce	9000000000	+a+aaaaaaa	agatataata	teteacceat	7140
tcctccaatg	tgtcagtcgc	geaegatgea	Letygeadaa	gggtgtacta	teccaccege	7200
gaccccacca	cccccttgc	gcgggctgcg	tgggagacag	ctagacacac	tecagteaat	
tcctggctag	gcaacatcat	catgtatgcg	cccaccttgt	gggcaaggat	gatcctgatg	7260
actcatttct	tctccatcct	totageteag	gaacaacttg	aaaaaqccct	agattgtcag	7320
	cctgttactc	anttanada	attasaatsa	ctcacatcat	tcaacgactc	7380
acctacgggg	cergilacie	Cattgageca	bb-b	-tt-	toddagaaat	7440
catggcctta	gcgcattttc	actccatagt	tactctccag	grgagarcaa	Lagggragger	
tcatqcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	7500
atecacacta	ggctactgtc	ccaaaaaaaa	aggactacca	cttqtqqcaa	gtacctcttc	7560
9	taaggaccaa	aataaaaata	actoraator	caactacate	ccagttggat	7620
aactgggcag	Laayyaccaa	guidadauci	accicaaccc		aatatataat	7680
ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	eergreregr	
qcccqacccc	gctggttcat	gtggtgccta	ctcctacttt	-ctgtaggggt	aggcatctat	7740
gtagtgggga	accgatgaac	ggggacctaa	acactccagg	ccaataggcc	atcctgtttt	7800
	ttttttttt	+++++++	++++++++	++++++++	tteteettt	7860
		اللانانانانانانانان		ntattn	tautuacco	7920
tttttcctct	ttttttcctt	LECTELECET	raaraaccc	accidagece	cagucacygu	
tagctgtgaa	aggtccgtga	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	7980
agatcaagta	ctcctgcagg	cgcgccacta	gtgggaatac	gcggggtatq	ccgcgtttta	8040
ggatattaa	gacccaattc	tcatchttca	cadcttatca	togataaget	ttaatacaat	8100
guararryac	gacccaaccc	ata'	and	~tat~a==t~	taadaatdda	8160
agtttatcac	agttaaattg	ctaacgcagt	caggcaccgt	grargaaarc	Laacaacycy	
ctcatcgtca	tcctcggcac	cgtcaccctg	gatgctgtag	gcataggctt	ggttatgccg	8220
atactaccaa	gcctcttgcg	ggatatcgtc	cattcccaca	gcatcgccaq	tcactatggc	8280
atactactac	cactatatac	attastacee	tttctatoco	cacccottct	cggagcactg	8340

tccgaccgct	ttggccgccg	cccagtcctg	ctcgcttcgc	tacttggagc	cactatcgac	8400
tacgcgatca	tggcgaccac	acccgtcctg	tggatcctct	acgccggacg	catcgtggcc	8460
ggcatcaccg	gcgccacagg	tgcggttgct	ggcgcctata	tcgccgacat	caccgatggg	8520
gaagatcggg	ctcgccactt	cgggctcatg	agcgcttgtt	tcggcgtggg	tatggtggca	8580
ggccccgtgg	ccgggggact	gttgggcgcc	atctccttgc	atgcaccatt	ccttgcggcg	8640
gcggtgctca	acggcctcaa	cctactactg	ggctgcttcc	taatgcagga	gtcgcataag	8700
ggagagcgtc	gaccgatgcc	cttgagagcc	ttcaacccag	tcagctcctt	ccggtgggcg	8760
	ctatcgtcgc					8820
caggtgccgg	cagcgctctg	ggtcattttc	ggcgaggacc	gctttcgctg	gagcgcgacg	8880
atgatcggcc	tgtcgcttgc	ggtattcgga	atcttgcacg	ccctcgctca	agccttcgtc	8940
actggtcccg	ccaccaaacg	tttcggcgag	aagcaggcca	ttatcgccgg	catggcggcc	9000
gacgcgctgg	gctacgtctt	gctggcgttc	gcgacgcgag	gctggatggc	cttccccatt	9060
atgattcttc	tcgcttccgg	cggcatcggg	atgcccgcgt	tgcaggccat	gctgtccagg	9120
	acgaccatca					9180
acttcgatca	ctggaccgct	gatcgtcacg	gcgatttatg	ccgcctcggc	gagcacatgg	9240
aacgggttgg	catggattgt	aggcgccgcc	ctataccttg	tetgeeteee	cgcgttgcgt	9300
cgcggtgcat	ggagccgggc	cacctcgacc	tgaatggaag	ccggcggcac	ctcgctaacg	9360
gattcaccac	tccaagaatt	ggagccaatc	aattettgeg	gagaactgtg	aatgcgcaaa	9420
ccaacccttg	gcagaacata	tccatcgcgt	ccgccatctc.	cagcagccgc	acgcggcgca	9480
	cgttgggtcc					9540
cccggctagg	ctggcggggt	tgccttactg	gttagcagaa	tgaatcaccg	atacgcgagc	9600
gaacgtgaag	cgactgctgc	tgcaaaacgt	ctgcgacctg	agcaacaaca	tgaatggtct	9660
teggttteeg	tgtttcgtaa	agtctggaaa	cgcggaagtc	agegeeetge	accarrargr	9720
tccggatctg	catcgcagga	tgctgctggc	taccctgtgg	aacacctaca	tetgtattaa	9780
cgaagcgctg	gcattgaccc	tgagtgattt	ttetetggte	ecgecgeate	cataccgcca'	9840
	ctcacaacgt					9900
	ctctcgtttc					9960 10020
acggaggcat	caagtgacca	aacaggaaaa	aaccgccctt	aacatggeee	getttateag	10020
aagccagaca	ttaacgcttc	tggagaaact	caacgagerg	gacgeggarg	tagagagttt	10140
	tcgcttcacg					10200
	ggtgaaaacc					10260
graageggar	gccgggagca	gacaageeeg	ccayggegeg	ctagegggtg	cattaaatat	10200
reggggegea	gccatgaccc	agreacycay	cyacaycyga	gratatasst	addadadada	10320
geggeareag	agcagattgt	actgagagtg	tattagastt	ggtgtgaaat	taaataaata	10380
rgcgraagga	gaaaataccg	carcaggege	taraatara	gaaagggggt	astaccetta	10500
tagaggag	ttcggctgcg caggggataa	gegageggta	aagatataaa	caaaggegge	aacacggcca	10560
aggaaggaat	aaaaggccgc	attactaca	ttttccata	andt conce	ccctaacaaa	10620
aggaaccgca	atcgacgctc	geegeegeeg	taacaaaaca	caacacact	ataaagatag	10680
anagagetta	ccctggaag	aagccagagg	cacteteeta	ttcccaccct	accacttacc	10740
gastagatat	ccgcctttct	acattagaa	agentagene	tttctcatac	ctcacactat	10800
aggataccigc	gttcggtgta	gatcattcac	tccaactcc	actatataca	cdaaccccc	10860
attraceca	accgctgcgc	cttatccoot	aactatcoto	ttgagtccaa	cccaataaaa	10920
geccageceg	cgccactggc	accaccogge	aaccaccgcc	ttaggagagg	gaggtatgta	10980
aacaatacta	cagagttctt	asataataa	cctaactacq	actacactac	aaggacagta	11040
	gegetetget					11100
tecqqeaac	aaaccaccgc	taataacaat	aatttttta	tttgcaagca	gcagattacg	11160
cacadeage	aaggatctca	agaagateet	ttgatctttt	ctacgggggtg	tgacgetcag	11220
tagaacaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaaa	gatetteace	11280
tagateettt	tctagataat	acqactcact	ata			11313
Lagaroore	JJUAJALUAL					

```
<210> 11
<211> 11184
<212> DNA
<213> Artificial Sequence
<220>
<223> Plasmid
```

```
gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180 gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgcccc 240
```

	t			++~+~~+		200
gcgagactgc	tagccgagta	grgrrgggre	gegaaaggee	Ligigglace	geergaragg	300
atacttacas.	gtgccccggg	aggtetegta	gaccatacac	catgaggagg	aatcctaaac	360
5-55-5			**********	taasttaasa	annanttata	420
Cicaaayaaa	aaccaaaggg	egegecatga	Ligaacaaga	cggaccgcac	geaggereee	
cggccgcttq	ggtggagagg	ctattcggct	atgactgggc	acaacagaca	atcggctgct	480
ctastaccac	cgtgttccgg	ctatcaacac	andudddddd	ggttcttttt	atcaagacca	540
cegacgeoge	292922229		~555575	2222222	5000050005	
acctgtccgg	tgccctgaat	gaactgcagg	acgaggcagc	geggetateg	rggerggeea	600
caacaaacat.	tccttgcgca	actatactca	acqttqtcac	tgaagcggga	agggactggc	660
-5 555-5-		2 2 2 2	taatataata	taraattaat	aataaaaaa	720
Lyctarrygg	cgaagtgccg	gggcaggacc	cccigccacc	ccaccicgci	cccgccgaga	
aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	gcttgatccg	gctacctgcc	780
cattccacca	ccaagcgaaa	categrateg	adcdaddadd	tactcoosto	gaaggggtg	840
caccogacca	ccaagegaaa	cacogcacog				
ttgtcgatca	ggatgatctg	gacgaagagc	arcagggger	egegeeagee	gaactgttcg	900
ccaggeteaa	ggcgcgcatg	cccaacaaca	aggatetegt	cataacccat	gacgatacct	960
~~++~~~~	Jo-J-J	annahaaaa	aattttataa	attaataaaa	tataaaaaa	1020
gerracegaa	tatcatggtg	gaaaatggcc	gettttttgg	acceategae	Lycygccygc	
tgggtgtggc	ggaccgctat	caggacatag	cgttggctac	ccgtgatatt	gctgaagagc	1080
ttaacaacaa	atgggctgac	cactteetea	tactttacaa	tategeeget	cccgattcgc	1140
ccggcggcga	acgggccgac				~~~~~~~~	
agcgcatcgc	cttctatcgc	cttcttgacg	agetettetg	agettaaaca	gaccacaacg	1200
atttccctct	agcgggatca	attccqcccc	tctccctccc	cccccctaa	cqttactqqc	1260
3	tggaataagg	aggatataga	tttatatata	tattattta	caccatatto	1320
egaageeget	Lygaalaayy	coggracaca	LLLYLCLALA	Lyclatic	caccacaccy	
ccgtcttttg	gcaatgtgag	ggcccggaaa	cctggccctg	tcttcttgac	gagcattcct	1380
aggggtgttt	cccctctcgc	caaaggaatg	caaggtetgt	tgaatgtcgt	gaaggaagga	1440
aggggcccc	CCCCCCCCC	caaaggaacg				1500
gttcctctgg	aagcttcttg	aagacaaaca	acgreegeag	egaecettig	caggcagcgg	
aaccccccac	ctggcgacag	atacetetae	qqccaaaaqc	cacgtgtata	agatacacct	1560
aaaaaaaaa	cacaacccca	ataaaaaatt	atasattaas	tanttatana	sacadteasa	1620
gcaaaggcgg	Cacaacccca	gigccacgii	grgagergga	ragicegega	aagagccaaa	
tggctctcct	caagcgtatt	caacaagggg	ctgaaggatg	cccagaaggt	accccattgt	1680
atgggatctg	atctggggcc	tcootocaca	tactttacat	atatttaatc	gaggttaaaa	1740
	~~~~~~~~~	33-3	+~~+++	±+000000	grataatard	1800
aacgictagg	cccccgaac	cacggggacg	Lygitticte	Lugadadduca	cgacaacacc	
atggcgccta	ttacggccta	ctcccaacag	acgcgaggcc	tacttggctg	catcatcact	1860
agggtgagag	gccgggacag	gaaccaggtc	aaaaaaaaaa	tccaaqtqqt	ctccaccqca	1920
	tcctggcgac	ataaataaat	aggatatatt	acactatata	testeatace	1980
acacaaccii	Lectiggegae	ctycytcaat	ggcgrgrgrr	ggaccgccca	ccacggegee	
ggctcaaaga	cccttgccgg	cccaaagggc	ccaatcaccc	aaatgtacac	caatgtggac	2040
caggaceteg	tcggctggca	addddddddd	agaacacatt	ccttgacacc	atqcacctqc	2100
	acctttactt	-2-2	astaccasta	tasttacaat	acaccaacaa	2160
ggcageregg	accitiacti	ggccacgagg	cacyccyacy	Leacedge	9090099099	
ggcgacagca	gggggagcct	actctccccc	aggcccgtct	cctacttgaa	gggetetteg	2220
aacaatccac	tgctctgccc	ctcqqqqcac	qctqtqqqca	tettteggge	tgccgtgtgc	2280
22-22	ttgcgaaggc	aataaaattt	atacccatca	antictationa	aaccactato	2340
accegagggg	Ligingaagge	ggrggacter	gracecycey	teresease		
cggtccccgg	tcttcacgga	caactcgtcc	cctccggccg	taccgcagac	attecaggig	2400
gcccatctac	acgcccctac	tggtagcggc	aagagcacta	aggtgccggc	tgcgtatgca	2460
adadasadat	ataaggtgct	tatactase	ccatccatca	acacacacat	aggittcggg	2520
gcccaagggc	acaaggegee	Lycockyaac	ccgcccgccg		~330000333	
gcgtatatgt	ctaaggcaca	tggtatcgac	cctaacatca	gaaccggggt	aaggaccatc	2580
accacaaata	ccccatcac	gtactccacc	tatqqcaaqt	ttcttqccga	cggtggttgc	2640
+a+aaaaaaa	cctatgacat	cataatatat	datasataca	actcaactca	ctcgaccact	2700
rccgggggg	CCLacgacac	cucuacucge	gacgagegee			
atcctgggca	tcggcacagt	cctggaccaa	gcggagacgg	ctggagcgcg	acceguegeg	2760
ctcaccacca	ctacgcctcc	gagatcagtc	accotoccac	atccaaacat	cqaqqaqqtq	2820
aatatataa	gcactggaga	aatacaattt	tataaraaaa	ccatccccat	coacaccatc	2880
gordragrada	gcaccggaga	aaccccccc	Lacygoddag	Coaroccac	ogagaooaco	
aaggggggga	ggcacctcat	tttctgccat	tccaagaaga	aatgtgatga	geregeegeg	2940
aaggtgtggg	gcctcggact	caatactata	gcatattacc	ggggccttga	tqtatccqtc	3000
	gcggagacgt	anttataata	aussauasuu 2	ctctaatcac	agactttacc	3060
ataccaacta	guggagaugu	carrigingua	gcaacggacg	Cucuacyac	999000000	
ggcgatttcg	actcagtgat	cgactgcaat	acatgtgtca	cccagacagt	cgacttcagc	3120
ctggacccga	ccttcaccat	taaaacaaca	accotoccac	aagacgcggt	gtcacgctcg	3180
	aaaaaataa	+24444444	atagagattt	acacette	gactroagga	3240
cageggegag	gcaggactgg	caggggcagg	augggcattt	acaygurugu	gaccccagga	
gaacggccct	cgggcatgtt	cgattcctcg	gttctgtgcg	agtgctatga	cgcgggctgt	3300
acttaataca	agctcacgcc	caccaaaacc	tcaattaaat	tacaaactta	cctaaacaca	3360
3000330003	agetetage	-3-3-5-5-	anattataaa	agagatatt	tagaggggtg	3420
ccagggctgc	ccgtctgcca	ggaccaccig	gageeeeggg	agagegeeee	Lacayyeee	
acccacatag	acgcccattt	cttgtcccag	actaagcagg	caggagacaa	cttcccctac	3480
ctootaocat	accaggctac	aatatacacc	agggctcagg	ctccacctcc	atcqtqqqac	3540
						3600
caaacycyga	agtgtctcat	acyguraday	cctacgctgc	acgagecaac	goodagaag	
tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accccataac	caaatacatc	3660
atagcataca	tgtcggctga	cctggaggtc	qtcacqaqca	cctgggtgct	ggtaggcgga	3720
atactacaa-	atatacaaa	mtattaaat-	2 2 2	acatactast	tataaaaaaa	3780
	ctctggccgc					
atcatcttgt	ccggaaagcc	ggccatcatt	cccgacaggg	aagtccttta	ccgggagttc	3840
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatco	aacaqqqaat	gcagctcqcc	3900
						3960
	aacagaaggc					
gctgctcccg	tggtggaatc	caagtggcgg	accctcgaag	ccttctgggc	gaagcatatg	4020
	tcagcgggat					4080
	cactgatggc					4140
						4200
accetectgt	ttaacatcct	gggggatgg	grggeegeee	aacttgctcc	Leccageget	4200

		• •				
acttetactt	tcatagacac	caacatcact	ggagcggctg	ttggcagcat	aggccttggg	4260
aarataatta	togotatttt	ggcaggttat	adsacsadaa	tggcaggcgc	actcataacc	4320
aaggrgerry	Lagararra	ggcaggccac	3303003333	tggttaacct	actecetact	4380
rccaaggcca	tgageggega	gatgetetee	accgaggace	angagatagt	acceccegee	4440
atcetetece	ctggcgccct	agregreggg	gregraraca	cagcgatact	gegeeggeae	
gtgggcccag	gggaggggc	tgtgcagtgg	atgaaccggc	tgatagcgtt	cgcttcgcgg	4500
ggtaaccacg	tctcccccac	gcactatgtg	cctgagagcg	acgctgcagc	acgtgtcact	4560
cagatcctct	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gtggatcaac	4620
gaggactgct	ccacaccata	ctccaactca	taactaaaaa	atgtttggga	ttggatatgc	4680
aggactgcc	atastttass	ascetacete	cadtccaadc	teetgeegeg	attaccagaa	4740
acggractga	tetantatas	gacceggeee	cagecoaage	addadadaa	ccccatcatc	4800
greceerrer	teteatglea	acgraggiac	aagggagtee	ggcggggcga	theestees	
caaaccacct	gcccatgtgg	agcacagate	acceggacary	tgaaaaacgg	Leccargagg	4860
atcgtggggc	ctaggacctg	tagtaacacg	tggcatggaa	cattccccat	taacgcgtac	4920
accacgggcc	cctgcacgcc	ctccccggcg	ccaaattatt	ctagggcgct	gtggcgggtg	4980
actactaaga	agtacgtgga	ggttacgcgg	gtgggggatt	tccactacgt	gacgggcatg	5040
accactgaca	acqtaaaqtq	cccatatcaa	attccaaccc	ccgaattctt	cacagaagtg	5100
astagaataa	aattacacaa	gtacgctcca	gcgtgcaaac	ccctcctacg	agagagatc	5160
gacggggcgc	tacacatass	tasstageta	attaaatcac	agctcccatg	caaacccaaa	5220
acattectgg	Logggotcaa	ttaatacttg	geegggeeae	agazasttag	cadacacaca	5280
ccggacgtag	cagigeteae	Liceargere	accyaccect	cccacattac	ggcggagacg	
gctaagcgta	ggctggccag	gggatetece	ccctccttgg	ccagctcatc	agetateeag	5340
ctgtctgcgc	cttccttgaa	ggcaacatgc	actacccgtc	atgactcccc	ggacgctgac	5400
ctcatcgagg	ccaacctcct	gtggcggcag	gagatgggcg	ggaacatcac	ccgcgtggag	5460
tcagaaaata	aggtagtaat	tttggagtct	ttcgagccgc	tccaagcgga	ggaggatgag	5520
agggaagtat	ccattccaac	ggagatcctg	cagagateca	ggaaattccc	tcgagcgatg	5580
agggaagcac	anataaaaa	ttacaaccct	ccactottao	agtcctggaa	agacccagac	5640
bb	cacteeegga	ccacaacccc	ttaccacata	ageoceggaa	taccatacca	5700
taegteeete	cagraguaca	cygytyteca	Ligitogettig	ccaaggcccc	teasttaaaa	5760
cctccacgga	ggaagaggac	ggttgtcctg	tcagaatcta	ccgtgtcttc	rgeerrggeg	
gagctcgcca	caaagacctt	cggcagctcc	gaatcgtcgg	ccgtcgacag	cggcacggca	5820
acggcctctc	ctggtgagga	cgtcgtctgc	tgctcgatgt	cctacacatg	gacaggcgcc	5880
ctgatcacgc	catgcgctgc	ggaggaaacc	aagctgccca	tcaatgcact	gagcaactct	5940
ttactccaac	accacaactt	ggtctatgct	acaacatctc	gcagcgcaag	cctgcggcag	6000
aadaaddtca	cctttgacag	actgcaggtc	ctagacgacc	actaccggga	catactcaaa	6060
aagaaggcca	agaagggatg	acceptana	actasacttc	tatccgtgga	agaagactat	6120
gagacgaagg	cgaaggcgcc	cacagecaag	gottaaacccc	atacacacasa	ggaageege	6180
aagctgacgc	cccacattc	ggeeagatet	aaatttggct	atggggcaaa	ggacgcccgg	6240
aacctatcca	gcaaggccgt	taaccacatc	cgctccgtgt	ggaaggactt	gerggaagae	
actgagacac	caattgacac	caccatcatg	gcaaaaaatg	aggttttctg	cgtccaacca	6300
gagaaggggg	gccgcaagcc	agctcgcctt	atcgtattcc	cagatttggg	ggttcgtgtg	6360
tgcgagaaaa	tggcccttta	cgatgtggtc	tccaccctcc	ctcaggccgt	gatgggctct	6420
tcatacggat	tccaatactc	tectagacag	caaatcaaat	tcctggtgaa	tqcctqqaaa	6480
acasassat	accetatora	cttcccatat	gadadddddt	gttttgactc	aacqqtcact	6540
gcgaagaaac	taaatattaa	accederace	taggaatgtt	gtgacttggc	ccccasacc	6600
gagaatgaca	Lecgiginga	ggagttaatt	caccaacycc	gegaceegge	ga	6660
agacaggcca	taaggteget	cacagagegg	CLLLacatog	ggggccccct	gaccaacccc	6720
aaagggcaga	actgcggcta	tegeeggtge	cgcgcgagcg	gtgtactgac	gaccagerge	
ggtaataccc	tcacatgtta	cttgaaggcc	gctgcggcct	gtcgagctgc	gaageteeag	6780
gactgcacga	tgctcgtatg	cggagacgac	cttgtcgtta	tctgtgaaag	cgcggggacc	6840
caaqaqqacq	aggcgagcct	acgggccttc	acggaggcta	tgactagata	ctctgccccc	6900
cctagagacc	cocccaaacc	agaatacgac	ttggagttga	taacatcatg	ctcctccaat	6960
atatasata	cacacaatac	atctgggaaa	agggtgtact	atctcacccg	tgaccccacc	7020
aggggggttg	agaaagataa	ataaaaaaa	actadacaca	ctccagtcaa	ttcctggcta	7080
acceceteg	tactatata	gegggagaea	goodgaoda	tastaatast	ractcatttc	7140
ggcaacacca	calguaty	geecacerry	cgggcaagga	tgatcctgat	gacccaccc	7200
ttctccatcc	ttctagctca	ggaacaactt	gaaaaagccc	tagattgtca	gatetacggg	
gcctgttact	ccattgagcc	acttgaccta	cctcagatca	ttcaacgact	ccatggcctt	7260
agcgcatttt	cactccatag	ttactctcca	ggtgagatca	atagggtggc	ttcatgcctc	7320
aggaaacttg	gggtaccgcc	cttgcgagtc	tggagacatc	gggccagaag	tgtccgcgct	7380
aggetactot	cccaaaaaaa	gagggctgcc	acttotogca	agtacctctt	caactgggca	7440
ataeaaacce	ageteaaact	cactccaatc	ccaactacat	cccagttgga	tttatccaqc	7500
taattaatta	ataattaaaa	addadaaaac	atatatcaca	gcctgtctcg	tacccasccc	7560
Lygilly	teteeteaa	cgggggagac	tatataccaca	taggatata	tatactaca	7620
egetggttea	tgtggtgcct	acticitatit	Lecgragggg	taggcatcta	ttttaccccc	7680
aaccgatgaa	cggggaccta	aacactccag	gccaataggc	catcctgttt	LLLLCCCCLLE	
				tttctccttt		7740
				ctagtcacgg		7800
aaggtccgtg	agccgcttga	ctgcagagag	tgctgatact	ggcctctctg	cagatcaagt	7860
actcctqcaq	gcgcgccact	agtgggaata	cgcggggtat	gccgcgtttt	agcatattga	7920
cgacccaatt	ctcatattta	acadettate	atcoataaoc	tttaatgcgg	tagtttatca	7980
cacttaaatt	actascacsa	trangracer	tatatasast	ctaacaatgc	actcatcatc	8040
						8100
				tggttatgcc		8160
ggcctcttgc	gggatategt	ccattccgac	agcatcgcca	gtcactatgg	cycyclycla	9700

	cgttgatgca					8220
	gcccagtcct					8280
	cacccgtcct					8340
	gtgcggttgc					8400
	tcgggctcat					8460
	tgttgggcgc					8520
	acctactact					8580
	ccttgagagc					8640
	ccgcacttat					8700
	gggtcatttt					8760
	cggtattcgg					8820
	gtttcggcga					8880
	tgctggcgtt					8940
	gcggcatcgg					9000
	agggacagct					9060
	tgatcgtcac					9120
	taggcgccgc					9180
	ccacctcgac					9240
	tggagccaat					9300
	atccatcgcg					9360
	ctggccacgg					9420
	ttgccttact					9480
	ctgcaaaacg					9540
	aagtctggaa					9600
	atgctgctgg					9660
	ctgagtgatt					9720
	ttccagtaac					9780
	catcggtatc					9840
	aaacaggaaa					9900
	ctggagaaac					9960
	gaccacgctg					10020
	ctctgacaca					10080
	agacaagccc					10140
agccatgacc	cagtcacgta	gcgatagcgg	agtgtatact	ggcttaacta	tgcggcatca	10200
	tactgagagt					10260
	gcatcaggcg					10320
	ggcgagcggt					10380
	acgcaggaaa					10440
	cgttgctggc					10500
	caagtcagag					10560
	gctccctcgt					10620
	tcccttcggg					10680
	aggtcgttcg					10740
	ccttatccgg					10800
	cagcagccac					10860
	tgaagtggtg					10920
	tgaagccagt					10980
	ctggtagcgg					11040
	aagaagatcc					11100
	aagggatttt		ttatcaaaaa	ggatcttcac	ctagatcctt	11160
ttctagataa	tacgactcac	tata				11184

```
<210> 12
<211> 11313
<212> DNA
<213> Artificial Sequence
<220>
<223> Plasmid
```

```
gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120 ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
```

gacgaccggg	teetttetta	gatcaacccg	ctcaatgcct	ggagatttgg	gagtgacaca	240
acasasctac	tagccgagta	atattaaatc	gcgaaaggcc	ttgtggtact	gcctgatagg	300
atacttacas	atacccaaa	aggtetegta	gaccgtgcac	catgaggagg	aatcctaaac	360
grgoregega	asaccasasa	cacaccataa	ttgaacaaga	tagattagac	acagattete	420
accadagada	aaccaaaggg	ctattccct	atgactgggc	acaacacaca	atcooctoct	480
atastagaaa	ggtggagagg	ctatcaggoe	aggggcgccc	aattetttt	atcaagacca	540
cigargeege	cycyccocy	gazatagaga	aggagagaaa	ggaaatataa	taactaacca	600
accigicegg	tgedetgaat	gaactgcagg	acgaggcagc	geggerareg	cggccggcca	660
cgacgggcgt	tccttgcgca	gergreereg	acgttgtcac	Lgaaguggga	agggactggc	
tgctattggg	cgaagtgccg	gggcaggatc	tcctgtcatc	teacettget	eetgeegaga	720
aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	gcttgatccg	gctacctgcc	780
cattcgacca	ccaagcgaaa	catcgcatcg	agcgagcacg	tactcggatg	gaagccggtc	840
ttgtcgatca	ggatgatctg	gacgaagagc	atcaggggct	cgcgccagcc	gaactgttcg	900
ccaggctcaa	ggcgcgcatg	cccgacggcg	aggatctcgt	cgtgacccat	ggcgatgcct	960
gcttgccgaa	tatcatggtg	gaaaatggcc	gcttttctgg	attcatcgac	tgtggccggc	1020
tgggtgtggc	ggaccgctat	caggacatag	cgttggctac	ccgtgatatt	gctgaagagc	1080
ttggcggcga	atgggctgac	cgcttcctcg	tgctttacgg	tatcgccgct	cccgattcgc	1140
agcgcatcgc	cttctatcgc	cttcttgacg	agttcttctg	agtttaaaca	gaccacaacg	1200
gtttccctct	agcgggatca	attecgeece	tetecetece	cccccctaa	cgttactggc	1260
caaaaccact	tggaataagg	ccqqtqtqcq	tttgtctata	tgttatttc	caccatattg	1320
ccatctttta	acaatataaa	qqcccqqaaa	catggaactg	tcttcttqac	gagcattcct	1380
aggggtcttt	cccctctcac	caaaqqaatq	caaggtctgt	tgaatgtcgt	gaaggaagca	1440
attectetaa	aaacttctta	aagagaaaga	acgtctgtag	cgaccctttg	cagacagcag	1500
ascacacas	ctaccacca	atacctetac	ggccaaaagc	cacatatata	agatacacct	1560
aacccccac	caggagaaaa	gtgccccgt	gtgagttgga	tacttataca	aadadtcaaa	1620
gcaaaggcgg	cacaacccca	graceacara	gtgagttgga	aggergegga	addageedda	1680
tggeteteet	caagegrace	taacaagggg	ctgaaggatg	ctctttagta	gaggttaaaa	1740
argggarerg	acecggggee	Leggtycaca	tgctttacat	ttanana	gaggicaaaa	1800
aacgtctagg	ccccccgaac	cacggggacg	tggttttcct	Ligadadada	cyaraaracc	
atggcgccta	ttacggccta	ctcccaacag	acgcgaggcc	tacttggetg	cateateaet	1860
agcctcacag	gccgggacag	gaaccaggtc	gagggggagg	tccaagtggt	ccccaccgca	1920
acacaatctt	tcctggcgac	ctgcgtcaat	ggcgtgtgtt	ggactgtcta	tcatggtgcc	1980
ggctcaaaga	cccttgccgg	cccaaagggc	ccaatcaccc	aaatgtacac	caatgtggac	2040
caggacctcg	tcggctggca	agcgcccccc	ggggcgcgtt	ccttgacacc	atgcacctgc	2100
ggcagctcgg	acctttactt	ggtcacgagg	catgccgatg	tcattccggt	gcgccggcgg	2160
ggcgacagca	gggggagcct	actctccccc	aggcccgtct	cctacttgaa	gggctcttcg	2220
ggcggtccac	tgctctgccc	ctcggggcac	gctgtgggca	tatttaggga	tgccgtgtgc	2280
acccgagggg	ttgcgaaggc	ggtggacttt	gtacccgtcg	agtctatgga	aaccactatg	2340
cgqtccccgg	tcttcacgga	caactcgtcc	cctccggccg	taccgcagac	attccaggtg	2400
gcccatctac	acqcccctac	tggtagcggc	aagagcacta	aggtgccggc	tgcgtatgca	2460
acccaagggt	ataaggtgct	tqtcctqaac	ccgtccgtcg	ccgccaccct	aggtttcggg	2520
			cctaacatca			2580
accacaaata	ccccatcac	gtactccacc	tatggcaagt	ttcttqccqa	cagtagttac	2640
			gatgagtgcc			2700
atectoooca	tradcaragt	cctogaccaa	gcggagacgg	ctagaacaca	actcatcata	2760
ctcaccacca	ctacacctcc	addatcaatc	accgtgccac	atccaaacat	caaaaaaata	2820
actatataca	acactacaca	aatccccttt	tatggcaaag	ccatccccat	cgagaccatc	2880
gerergreea	gcaccggaga	tttataaat	tccaagaaga	aatotoooda	acteaceaca	2940
aagggggga	ggcaccccac	asstratate	gcatattacc	aaagaaattaa	tatatecate	3000
addetgeeeg	geereggaer	catteteeta	gcaacggacg	chchaatcac	agaatttaca	3060
						3120
			acatgtgtca			3180
ctggacccga	cctcaccat	tgagacgacg	accgtgccac	aagacgcggt	greategereg	3240
cagcggcgag	gcaggactgg	taggggcagg	atgggcattt	acaggirigi	gaeteeagga	
			gttctgtgcg			3300
			tcagttaggt			3360
ccagggttgc	ccgtctgcca	ggaccatctg	gagttctggg	agagcgtctt	tacaggcctc	3420
acccacatag	acgcccattt	cttgtcccag	actaagcagg	caggagacaa	cttcccctac	3480
			agggctcagg			3540
			cctacgctgc			3600
tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accccataac	caaatacatc	3660
atggcatgca	tgtcagctga	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	3720
gtcctagcag	ctctggccgc	gtattgcctg	acaacaggca	gcgtggtcat	tgtgggcagg	3780
			cccgacaggg			3840
			ccttacatcg			3900
			ctgcaaacag			3960
			accctcgaag			4020
			gcaggcttgt			4080
			tctatcacca			4140
J-JJJ				JJ		

accetectot	ttaacatcct	ggggggatgg	gtggccgccc	aacttqctcc	tcccagcqct	4200
			ggagcggctg			4260
geceeegeee	t	aggaagttat	227262262	taggeageae	actactacaa	4320
			ggagcagggg			
			accgaggacc			4380
atcctctccc	ctggcgccct	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4440
			atgaaccggc			4500
			cctgagagcg			4560
			ctgctgaaga			4620
			tggctaagag			4680
			cagtccaagc			4740
			aagggagtct			4800
caaaccacct	gcccatgtgg	agcacagatc	accggacatg	tgaaaaacgg	ttccatgagg	4860
			tggcatggaa			4920
			ccaaattatt			4980
			gtgggggatt			5040
						5100
			gttccggccc			
			gcgtgcaaac			5160
			gttgggtcac			5220
ccggacgtag	cagtgctcac	ttccatgctc	accgacccct	cccacattac	ggcggagacg	5280
gctaagcgta	ggctggccag	gggatctccc	ccctccttgg	ccagctcatc	agctagccag	5340
			actacccgtc			5400
			gagatgggcg			5460
						5520
			ttcgagccgc			
			cggaggtcca			5580
			ccactgttag			5640
tacgtccctc	cagtggtaca	cgggtgtcca	ttgccgcctg	ccaaggcccc	tccgatacca	5700
cctccacgga	ggaagaggac	ggttgtcctg	tcagaatcta	ccgtgtcttc	tgccttggcg	5760
			gaatcgtcgg			5820
			ggcgacgcgg			5880
			ggggatcccg			5940
			gtcgtctgct			6000
			gaggaaacca			6060
			gtctatgcta			6120
ctgcggcaga	agaaggtcac	ctttgacaga	ctgcaggtcc	tggacgacca	ctaccgggac	6180
			acagttaagg			6240
gaagcctgta	agctgacgcc	cccacattcq	gccagatcta	aatttggcta	tggggcaaag	6300
			aaccacatcc			6360
			accatcatgg			6420
ctggaagaca	ccgagacacc	aaccgacacc	accaccacgg	taatattaaa	agatttaaga	6480
			gctcgcctta			
			gatgtggtct			6540
atgggctctt	catacggatt	ccaatactct	cctggacagc	gggtcgagtt	cctggtgaat	6600
			ttcgcatatg			6660
acggtcactg	agaatgacat	ccgtgttgag	gagtcaatct	accaatgttg	tgacttggcc	6720
cccgaagcca	gacaggccat	aaggtcgctc	acagagcggc	tttacatcgg	agaccccta	6780
			cgccggtgcc			6840
			ttgaaggccg			6900
						6960
			ggagacgacc			
			cgggccttca			7020
			gaatacgact			7080
tcctccaatg	tgtcagtcgc	gcacgatgca	tctggcaaaa	gggtgtacta	tctcacccgt	7140
gaccccacca	acacacttga	gcgggctgcg	tgggagacag	ctagacacac	tccagtcaat	7200
			cccaccttgt			7260
			gaacaacttg			7320
			cttgacctac			7380
						7440
			tactctccag			
			ttgcgagtct			7500
			agggctgcca			7560
aactgggcag	taaggaccaa	gctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	7620
ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	7680
			ctcctacttt			7740
			acactccagg			7800
			tttttttt			7860
						7920
			tggtggctcc			
			tgcagagagt			7980
			gtgggaatac			8040
gcatattgac	gacccaattc	tcatgtttga	cagcttatca	tcgataagct	ttaatgcggt	8100

agtttatcac	agttaaattg	ctaacgcagt	caggcaccgt	gtatgaaatc	taacaatgcg	8160
ctcatcqtca	tcctcqgcac	cgtcaccctg	gatgctgtag	gcataggctt	ggttatgccg	8220
gtactgccgg	gcctcttgcg	ggatatcgtc	cattccgaca	gcatcgccag	tcactatggc	8280
			tttctatgcg			8340
			ctcgcttcgc			8400
tacgcgatca	tggcgaccac	accegteetg	tggatcctct	acgccggacg	categtggee	8460
ggcatcaccg	gcgccacagg	tgcggttgct	ggcgcctata	tcgccgacat	caccgatggg	8520
gaagatcggg	ctcaccactt	caaactcata	agcgcttgtt	tcaacataaa	tatggtggca	8580
			atctccttgc			8640
acaatactca	acggcctcaa	cctactactq	ggctgcttcc	taatqcaqqa	gtcgcataag	8700
adadadcatc	gaccgatgcc	cttgagagcc	ttcaacccag	tcagctcctt	ccaataaaca	8760
caaaacataa	chatcotcoc	cacacttata	actgtcttct	ttatcatgca	actcotagoa	8820
			ggcgaggacc			8880
atgatcggc	tateacttac	ggtattcgga	atcttgcacg	ccctcactca	agectteate	8940
			aagcaggcca			9000
			gcgacgcgag			9060
atrattette	tracttraa	caacatcaa	atgcccgcgt	tacadaccat	actatccaaa	9120
cagattacata	accaccatca	addacadatt	caaggatcgc	tcacaactct	taccaccta	9180
			gcgatttatg			9240
			ctataccttg			9300
			tgaatggaag			9360
cycygrycar	tagaagaatt	ggaggaatg	aattcttgcg	ceggeggeae	aatacacaaa	9420
						9480
			ccgccatctc			9540
			tgcgcatgat			9600
ceeggeragg	erggeggggt	tagectacty	gttagcagaa	rgaarcaccy	tacted	9660
gaacgtgaag	egaetgetge	tgcaaaacgt	ctgcgacctg	agcaacaaca	rgaarggree	
			cgcggaagtc			9720
			taccctgtgg			9780
cgaagcgctg	gcattgaccc	tgagtgattt	ttctctggtc	ccgccgcatc	catacegeca	9840
			gggcatgttc			9900
			ttacccccat			9960
acggaggcat	caagtgacca	aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	10020
			caacgagctg			10080
			tgagctttac			10140
			gcagctcccg			10200
			tcagggcgcg			10260
			cgatagcgga			10320
			caccatatgc			10380
tgcgtaagga	gaaaataccg	catcaggcgc	tetteegett	cctcgctcac	tgactcgctg	10440
			tcagctcact			10500
tccacagaat	caggggataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	10560
aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	10680
caggcgtttc	cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	10740
			agcgtggcgc			10800
			tccaagctgg			10860
			aactatcgtc			10920
			ggtaacagga			10980
			cctaactacg			11040
tttggtatct	gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	11100
tccggcaaac	aaaccaccgc	tggtagcggt	ggtttttttg	tttgcaagca	gcagattacg	11160
			ttgatctttt			11220
tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	11280
	tctagataat					11313

```
<210> 13
<211> 11184
<212> DNA
<213> Artificial Sequence
<220>
<223> Plasmid
```

gccagccccc	gattgggggc	gacactccac	catagatcac	tcccctqtqa	ggaactactg	60
		agccatggcg				120
		tagtggtctg				180
		gatcaacccg				240
gcgagactgc	tagccgagta	gtgttgggtc	gcgaaaggcc	ttgtggtact	gcctgatagg	300
		aggtctcgta				360
		cgcgccatga				420
		ctattcggct				480
		ctgtcagcgc				540
acctgtccgg	tgccctgaat	gaactgcagg	acgaggcagc	gcggctatcg	tggctggcca	600
cgacgggcgt	tccttgcgca	gctgtgctcg	acgttgtcac	tgaagcggga	agggactggc	660
		gggcaggatc				720
		gcaatgcggc				780
		catcgcatcg				840
ttgtcgatca	ggatgatctg	gacgaagagc	atcaggggct	cgcgccagcc	gaactgttcg	900
ccaggctcaa	ggcgcgcatg	cccgacggcg	aggatctcgt	cgtgacccat	ggcgatgcct	960
		gaaaatggcc				1020
		caggacatag				1080
						1140
		cgcttcctcg				
		cttcttgacg				1200
gtttccctct	agcgggatca	attccgcccc	tetecetece	cccccctaa	cgttactggc	1260
cgaagccgct	tggaataagg	ccggtgtgcg	tttgtctata	tgttattttc	caccatattg	1320
		ggcccggaaa				1380
		caaaggaatg				1440
		aagacaaaca				1500
aaccccccac	ctggcgacag	gtgcctctgc	ggccaaaagc	cacgtgtata	agatacacct	1560
gcaaaggcgg	cacaacccca	gtgccacgtt	gtgagttgga	tagttgtgga	aagagtcaaa	1620
		caacaagggg				1680
		tcggtgcaca				1740
						1800
		cacggggacg				
atggcgccta	ttacggccta	ctcccaacag	acgcgaggcc	tacttggctg	catcatcact	1860
agcctcacag	gccgggacag	gaaccaggtc	gagggggagg	tccaagtggt	ctccaccgca	1920
acacaatctt	tcctggcgac	ctgcgtcaat	ggcgtgtgtt	ggactgtcta	tcatggtgcc	1980
		cccaaagggc				2040
		agegeeece				2100
						2160
		ggtcacgagg				
		actctccccc				2220
ggcggtccac	tgctctgccc	ctcggggcac	gctgtgggca	tctttcgggc	tgccgtgtgc	2280
acccgaggg	ttgcgaaggc	ggtggacttt	gtacccgtcg	agtctatgga	aaccactatg	2340
		caactcgtcc				2400
		tggtagcggc				2460
		tgtcctgaac				2520
		tggtatcgac				2580
accacgggtg	cccccatcac	gtactccacc	tatggcaagt	ttcttgccga	cggtggttgc	2640
		cataatatgt				2700
		cctggaccaa				2760
						2820
		gggatcggtc				
		aatccccttt				2880
		tttctgccat				2940
aagctgtccg	gcctcggact	caatgctgta	gcatattacc	ggggccttga	tgtatccgtc	3000
		cattgtcgta				3060
		cgactgcaat				3120
						3180
		tgagacgacg				
		taggggcagg				3240
gaacggccct	cgggcatgtt	cgattcctcg	gttctgtgcg	agtgctatga	cgcgggctgt	3300
gcttggtacg	agctcacgcc	cgccgagacc	tcagttaggt	tgcgggctta	cctaaacaca	3360
		ggaccatctg				3420
		cttgtcccag				3480
		ggtgtgcgcc				3540
		acggctaaag				3600
		aaacgaggtt				3660
		cctggaggtc				3720
		gtattgcctg				3780
		ggccatcatt				3840
		ctcacacctc				3900
gaacatttca	aacagaaggc	aatcgggttg	ctgcaaacag	ccaccaagca	agcggaggct	3960
			<del>-</del>			

	actactacca	cqqtqqaatc	caagtggcgg	accetegaag	ccttctgggc	gaagcatatq	4020
	tqqaatttca	tcaqcqqqat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4080
			attcacagcc				4140
	accetectet	ttaacatcct	ggggggatgg	gtggccgccc	aacttgctcc	tcccagcgct	4200
	gcttctgctt	tcgtaggcgc	cggcatcgct	ggagcggctg	ttggcagcat	aggccttggg	4260
			ggcaggttat				4320
			gatgccctcc				4380
	atcctctccc	ctagcaccct	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4440
	atagacccag	aaaaaaaaaa	tgtgcagtgg	atgaaccggc	tgatagcgtt	cgcttcqcqq	4500
			gcactatgtg				4560
			catcactcag				4620
			ctccggctcg				4680
			gacctggctc				4740
	qtccccttct	tctcatqtca	acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	4800
			agcacagatc				4860
			tagtaacacg				4920
			ctccccggcg				4980
	gctgctgagg	aqtacqtqqa	ggttacgcgg	gtgggggatt	tccactacgt	gacgggcatg	5040
	accactgaca	acqtaaaqtq	cccgtgtcag	gttccggccc	ccgaattctt	cacagaagtg	5100
	gatggggtgc	gattacacaq	gtacgctcca	gcgtgcaaac	ccctcctacg	ggaggaggtc	5160
	acattcctqq	tcaaactcaa	tcaatacctg	gttgggtcac	agctcccatg	cgagcccgaa	5220
	ccaacataa	cagtgctcac	ttccatgctc	accgacccct	cccacattac	ggcggagacg	5280
	actaaacata	gactagccag	gggatctccc	ccctccttqq	ccaqctcatc	agctatccag	5340
	ctatctacac	cttccttgaa	ggcaacatgc	actacccqtc	atgactcccc	ggacgctgac	5400
	ctcatcgagg	ccaacctcct	gtggcggcag	gagatggggg	qqaacatcac	ccacataaaa	5460
	tcagaaaata	aggtagtaat	tttggagtct	ttcaaaccac	tccaaqcqqa	qqaqqatqaq	5520
	agggaagtat	ccattccaac	ggagatcctg	cagaggtcca	ggaaattccc	tcaaacaata	5580
	cccatatogo	cactcccaga	ttacaaccct	ccactgttag	agtcctggaa	ggacccggac	5640
			cgggtgtcca				5700
			ggttgtcctg				5760
	gagetegea	caaagaggtt	cggcagctcc	gaat.cgt.cgg	ccatcaacaa	caacacaaca	5820
	accocctctc	ctaataaaaa	cgtcgtctgc	tactcaatat	cctacacato	gacaggggg	5880
	ctgatcacgc	catgcgctgc	ggaggaaacc	aagctgccca	tcaatgcact	gagcaactct	5940
	ttgctccgac	accacaactt	ggtctatgct	acaacatctc	acaacacaaa	cctacaacaa	6000
	aagaaggtca	cctttgacag	actgcaggtc	ctggacgacc	actaccogga	catactcaaa	6060
-	gagatgaagg	caaaaacatc	cacagttaag	gctaaacttc	tatccqtqqa	ggaagcctgt	6120
	aagctgacgc	cccacattc	ggccagatct	aaatttggct	atggggcaaa	ggacgtccgg	6180
	aacctatcca	acaaaaccat	taaccacatc	cactccatat	qqaaqqactt	gctggaagac	6240
	actgagacac	caattgacac	caccatcatg	gcaaaaaatg	aggttttctg	cqtccaacca	6300
	gagaaggggg	accacaaacc	agctcgcctt	atcqtattcc	cagatttqqq	ggttcgtgtg	6360
	tacaagaaaa	taaccettta	cgatgtggtc	tccaccctcc	ctcaggccgt	gatgggctct	6420
	tcatacqqat	tccaatactc	tcctggacag	cagatcaaat	tcctqqtqaa	tgcctggaaa	6480
	gcgaagaaat	accctataga	cttcgcatat	gacacccgct	qttttqactc	aacggtcact	6540
			ggagtcaatc				6600
	agacaggca	taaggtcgct	cacagagcgg	ctttacatcq	gagaccccct	gactaattct	6660
	aaaqqqcaqa	actocoocta	tcgccggtgc	cqcqcqaqcq	gtgtactgac	gaccagctgc	6720
	ggtaataccc	tcacatotta	cttgaaggcc	gctgcggcct	gtcgagctgc	gaagctccag	6780
	gactgcacga	tactcatata	cggagacgac	cttgtcgtta	tctgtgaaag	cgcggggacc	6840
			acgggccttc				6900
	cctggggacc	cqcccaaacc	agaatacgac	ttggagttga	taacatcatg	ctcctccaat	6960
	atatcaatca	cacacaatac	atctggcaaa	agggtgtact	atctcacccg	tgaccccacc	7020
	acccccttq	cacaaactac	gtgggagaca	qctaqacaca	ctccaqtcaa	ttcctggcta	7080
	ggcaacatca	tcatatatac	gcccaccttg	tagacaagga	tgatcctgat	gactcatttc	7140
	ttctccatcc	ttctagctca	ggaacaactt	gaaaaagccc	tagattgtca	gatctacqqq	7200
			acttgaccta				7260
	aggggatttt	cactccatag	ttactctcca	gataagatca	atagggtggc	ttcatqcctc	7320
			cttgcgagtc				7380
			gagggctgcc				7440
			cactccaatc				7500
			cgggggagac				7560
			actcctactt				7620
			aacactccag				7680
	ttttttttt	tttttttt	tttttttt	tttttttt	tttctccttt	ttttttcctc	7740
			ttggtggctc				7800
			ctgcagagag				7860
	actcctqcaq	gcgcaccact	agtgggaata	cacaaaatat	gccacatttt	agcatattga	7920
	J J			5 5555	5 5 5 5 5 5 5		

coacceatt	ctcatattta	acagettate	atcoataago	tttaatgcgg	tagtttatca	7980
				ctaacaatgc		8040
						8100
				tggttatgcc		
				gtcactatgg		8160
				tcggagcact		8220
				ccactatcga		8280
				gcatcgtggc		8340
ggcgccacag	gtgcggttgc	tggcgcctat	atcgccgaca	tcaccgatgg	ggaagatcgg	8400
gctcgccact	tcgggctcat	gagcgcttgt	ttcggcgtgg	gtatggtggc	aggccccgtg	8460
gccgggggac	tgttgggcgc	catctccttg	catgcaccat	tccttgcggc	ggcggtgctc	8520
aacggcctca	acctactact	gggctgcttc	ctaatgcagg	agtcgcataa	gggagagcgt	8580
				tccggtgggc		8640
				aactcgtagg		8700
				ggagcgcgac		8760
				aagccttcgt		8820
				gcatggcggc		8880
				ccttccccat		8940
						9000
				tgctgtccag		9060
				ttaccagcct		9120
				cgagcacatg		
				ccgcgttgcg		9180
				cctcgctaac		9240
				gaatgcgcaa		9300
				cacgcggcgc		9360
gcgttgggtc	ctggccacgg	gtgcgcatga	tegtgeteet	gtcgttgagg	acccggctag	9420
gctggcgggg	ttgccttact	ggttagcaga	atgaatcacc	gatacgcgag	cgaacgtgaa	9480
gcgactgctg	ctgcaaaacg	tctgcgacct	gagcaacaac	atgaatggtc	ttcggtttcc	9540
gtgtttcgta	aagtctggaa	acgcggaagt	cagcgccctg	caccattatg	ttccggatct	9600
				atctgtatta		9660
				ccataccgcc		9720
				aacccgtatc		9780
				ttccccctta		9840
				cgctttatca		9900
				gaacaggcag		9960
				ctcgcgcgtt		10020
						10080
				acagettgte		10140
				gttggcgggt		10200
				ggcttaacta		
				taccgcacag		10260
agaaaatacc	gcatcaggcg	ctcttccgct	tcctcgctca	ctgactcgct	gegeteggte	10320
				taatacggtt		10380
tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaccgt	10440
aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatcacaaa	10500
				tataaagata		10560
cccctggaa	gctccctcgt	gcgctctcct	gttccgaccc	tgccgcttac	cggatacctg	10620:
teegeettte	teeetteggg	aagcgtggcg	ctttctcata	gctcacgctg	taggtatctc	10680
				acgaaccccc		10740
				acccggtaag		10800
tcaccactaa	cagcagccac	tagtaacagg	attagcagag	cgaggtatgt	aggcggtgct	10860
				gaaggacagt		10920
				gtagctcttg		10980
				agcagattac		11040
				ctgacgctca		11100
				ggatcttcac		11160
			LLALUAAAAA	ggallillat	clagacticic	11184
Liciayataa	tacgactcac	caca				*****

```
<210> 14
<211> 11184
<212> DNA
<213> Artificial Sequence
<220>
<223> Plasmid
```

<#U	0 > T#					
gccagccccc	gattgggggc	gacactccac	catagatcac	teceetgtga	ggaactactg	60
tcttcacgca	gaaagcgtct	agccatggcg	ttagtatgag	tgtcgtgcag	cctccaggac	120
					gaattgccag	180
gacgaccggg	teetttettg	gatcaacccg	ctcaatgcct	ggagatttgg	gegtgeecee	240
gcgagactgc	tagccgagta	gtgttgggtc	gcgaaaggcc	ttgtggtact	qcctqataqq	300
gtgcttgcga	gtgccccggg	aggtctcgta	gaccgtgcac	catgagcacg	aatcctaaac	360
ctcaaagaaa	aaccaaaggg	cgcgccatga	ttgaacaaga	tggattgcac	qcaqqttctc	420
cggccgcttg	ggtggagagg	ctattcggct	atgactgggc	acaacagaca	atcaactact	480
ctgatgccgc	catattccaa	ctqtcaqcqc	agggggggc	ggttcttttt	atcaagacca	540
acctgtccgg	tgccctgaat	gaactgcagg	acqaqqcaqc	gcggctatcg	taactaacca	600
				tgaagcggga		660
				tcaccttgct		720
				gcttgatccg		780
				tactcggatg		840
				cgcgccagec		900
ccaggctcaa	gacacacata	cccaacaca	aggatetegt	cgtgacccat	gacastacet	960
acttaccaaa	tatcatooto	gaaaatggcc	acttttctaa	attcatcgac	tataaccaac	1020
				ccgtgatatt		1080
				tatcgccgct		1140
				agtttaaaca		1200
				cccccctaa		1260
				tgttattttc		1320
				tcttcttgac		1380
				tgaatgtcgt		1440
				cgaccctttg		1500
				cacgtgtata		1560
acassaacaa	cacaacccca	ataccacatt	ataaattaa	tagttgtgga	agacacaccc	1620
taactctcct	caagggtatt	gegeeaegee	ctaaacata	cccagaaggt	aagagccaaa	1680
atoggateto	atctgggacc	teactagggg	tactttacat	gtgtttagtc	accccarrge	1740
				ttgaaaaaca		1800
				tacttggctg		1860
				tccaagtggt		1920
				ggactgtcta		1980
				aaatgtacac		2040
				ccttgacacc		2100
				tcattccggt		2160
				cctacttgaa		2220
				tettteggge		2280
				agtctatggg		2340
cggcccccgg	ccccacgga	basecegeee	eereeggeeg	taccgcagac	actecaggtg	2400
geceatetae	acgecectae	tggtagegge	aagagcacta	aggtgccggc	tgcgtatgca	2460
geceaagggt	ataaggtget	tgteetgaae	cegreegreg	ccgccaccct	aggtttcggg	2520
				gaaccggggt		2580
				ttcttgccga		2640
				actcaactga		2700
				ctggagcgcg		2760
cregecaceg	ctacgcctcc	gggatcggtc	accgtgccac	atccaaacat	cgaggaggtg	2820
getetgteea	gcactggaga	aatccccttt	tatggcaaag	ccatccccat	cgagaccatc	2880
aagggggga	ggcaccccat	tttctgccat	tccaagaaga	aatgtgatga	geregeegeg	2940
aagctgtccg	gcctcggact	caatgctgta	gcatattacc	ggggccttga	tgtatccgtc	3000
ataccaacta	gcggagacgt	cattgtcgta	gcaacggacg	ctctaatgac	gggctttacc	3060
ggcgatttcg	actcagtgat	cgactgcaat	acatgtgtca	cccagacagt	cgacttcagc	3120
ctggacccga	ccttcaccat	tgagacgacg	accgtgccac	aagacgcggt	gtcacgctcg	3180
cagcggcgag	gcaggactgg	taggggcagg	atgggcattt	acaggtttgt	gactccagga	3240
gaacggccct	cgggcatgtt	cgattcctcg	gttctgtgcg	agtgctatga	cgcgggctgt	3300
gcttggtacg	agctcacgcc	cgccgagacc	tcagttaggt	tgcgggctta	cctaaacaca	3360
ccagggttgc	ccgtctgcca	ggaccatctg	gagttctggg	agagcgtctt	tacaggcctc	3420
acccacatag	acgcccattt	cttgtcccag	actaagcagg	caggagacaa	cttcccctac	3480
crggragcat	accaggctac	ggrgrgcgcc	agggctcagg	ctccacctcc	atcgtgggac	3540
caaatgtgga	agtgtctcat	acggctaaag	cctacgctgc	acgggccaac	gecectgetg	3600
tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accccataac	caaatacatc	3660
atggcatgca	tgtcagctga	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	3720
gtcctagcag	ctctggccgc	gtattgcctg	acaacaggca	gcgtggtcat	tgtgggcagg	3780
accatcttgt	ccggaaagcc	ggccatcatt	cccgacaggg	aagtctttta	ccgggagttc	3840
gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacggggaat	gcagctcgcc	3900

gaacatttca	aacagaaggc	aatcgggttg	ctgcaaacag	ccaccaagca	agcqqaqqct	3960
				ccttctgggc		4020
				ccactctgcc		4080
				gcccgctcac		4140
				aacttgctcc		4200
gerrergerr	regraggege	eggeateget	ggageggetg	ttggcagcat	aggccttggg	4260
				. tggcaggcgc		4320
				tggttaacct		4380
atcctctccc	ctggcgccct	agtcgtcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	4440
				tgatagcgtt		4500
				acgctgcagc		4560
				ggcttcacca		4620
				atgtttggga		4680
				tcctgccgcg		4740
atacattat	tatastatas	acceggeee	agecouage	accedecaca	accyccygya	
				ggcggggcga		4800
				tgaaaaacgg		4860
				cattccccat		4920
				ctagggcgct		4980
				tccactacgt		5040
accactgaca	acgtaaagtg	cccgtgtcag	gttccggccc	ccgaattctt	cacagaagtg	5100
				ccctcctacg		5160
				agctcccatg		5220
				cccacattac		5280
				ccagctcatc		5340
				atgactcccc		5400
						5460
				ggaacatcac		
				tccaagcgga		5520
				ggaaattccc		5580
				agtcctggaa		5640
				ccaaggcccc		5700
				ccgtgtcttc		5760
gagctcgcca	caaagacctt	cggcagctcc	gaatcgtcgg	ccgtcgacag	cggcacggca	5820
acggcctctc	ctggtgagga	cgtcgtctgc	tgctcgatgt	cctacacatg	gacaggcgcc	5880
ctgatcacgc	catgcgctgc	ggaggaaacc	aagctgccca	tcaatgcact	gagcaactct	5940
				gcagcgcaag		6000
				actaccggga		6060
				tatccgtgga		6120
				atggggcaaa		6180
				ggaaggactt		6240
				aggttttctg		6300
						6360
				cagatttggg		
				ctcaggccgt		6420
				tcctggtgaa		6480
				gttttgactc		6540
				gtgacttggc		6600
				ggggccccct		6660
				gtgtactgac		6720
ggtaataccc	tcacatgtta	cttgaaggcc	gctgcggcct	gtcgagctgc	gaagctccag	6780
gactgcacga	tgctcgtatg	cggagacgac	cttgtcgtta	tctgtgaaag	cgcggggacc	6840
caagaggacg	aggcgagcct	acgggccttc	acggaggcta	tgactagata	ctctgcccc	6900
				taacatcatg		6960
				atctcacccg		7020
				ctccagtcaa		7080
				tgatcctgat		7140
				tagattgtca		7200
						7260
goodgecace	ccattgagee	accegaceca	ccccagacca	ttcaacgact	thankanaka	
				atagggtggc		7320
				gggccagaag		7380
aggctactgt	cccagggggg	gagggctgcc	acttgtggca	agtacctctt	caactgggca	7440
graaggacca	agctcaaact	cactccaatc	ccggctgcgt	cccagttgga	tttatccagc	7500
tggttcgttg	ctggttacag	cgggggagac	atatatcaca	gcctgtctcg	tgcccgaccc	7560
cgctggttca	tgtggtgcct	actcctactt	tctgtagggg	taggcatcta	tctactcccc	7620
aaccgatgaa	cggggaccta	aacactccag	gccaataggc	catcctgttt	ttttcccttt	7680
tttttttct	tttttttt	ttttttttt	ttttttttt	tttctccttt	ttttttcctc	7740
tttttttcct	tttctttcct	ttggtggctc	catcttagcc	ctagtcacgg	ctagctgtga	7800
aaggtccgtq	agccgcttga	ctgcagagag	tgctgatact	ggcctctctg	cagatcaagt	7860

actcctgca	g gcgcgccact	agtgggaata	cgcggggtat	gccgcgtttt	agcatattga	7920
	t ctcatgtttg					7980
cagttaaat	t gctaacgcag	tcaggcaccg	tgtatgaaat	ctaacaatgc	gctcatcgtc	8040
atcctcgg	ca ccgtcaccct	ggatgctgta	ggcataggct	tggttatgcc	ggtactgccg	8100
ggcctcttg	gc gggatatcgt	ccattccgac	agcatcgcca	gtcactatgg	cgtgctgcta	8160
gcgctatat	g cgttgatgca	atttctatgc	gcacccgttc	tcggagcact	gtccgaccgc	8220
tttggccg	cc gcccagtcct	gctcgcttcg	ctacttggag	ccactatcga	ctacgcgatc	8280
atggcgaco	ca cacccgtcct	gtggatcctc	tacgccggac	gcatcgtggc	cggcatcacc	8340
ggcgccaca	ag gtgcggttgc	tggcgcctat	atcgccgaca	tcaccgatgg	ggaagatcgg	8400
gctcgccad	ct tegggeteat	gagcgcttgt	ttcggcgtgg	gtatggtggc	aggccccgtg	8460
gccggggg	ac tgttgggcgc	catctccttg	catgcaccat	tccttgcggc	ggcggtgctc	8520
aacggccto	ca acctactact	gggctgcttc	ctaatgcagg	agtcgcataa	gggagagcgt	8580
cgaccgato	gc ccttgagagc	cttcaaccca	gtcagctcct	tccggtgggc	gcggggcatg	8640
actatcgto	eg cegeacttat	gactgtcttc	tttatcatgc	aactcgtagg	acaggtgccg	8700
gcagcgctc	t gggtcatttt	cggcgaggac	cgctttcgct	ggagcgcgac	gatgatcggc	8760
ctgtcgctt	g cggtattcgg	aatcttgcac	gccctcgctc	aagccttcgt	cactggtccc	8820
gccaccaaa	ıc gtttcggcga	gaagcaggcc	attatcgccg	gcatggcggc	cgacgcgctg	8880
ggctacgto	et tgctggcgtt	cgcgacgcga	ggctggatgg	ccttccccat	tatgattctt	8940
ctcgcttcc	g gcggcatcgg	gatgcccgcg	ttgcaggcca	tgctgtccag	gcaggtagat	9000
gacgaccat	c agggacagct	tcaaggatcg	ctcgcggctc	ttaccagcct	aacttcgatc	9060
actggaccg	jc tgatcgtcac	ggcgatttat	gccgcctcgg	cgagcacatg	gaacgggttg	9120
gcatggatt	g taggcgccgc	cctatacctt	gtctgcctcc	ccgcgttgcg	tcgcggtgca	9180
tggagccgg	g ccacctcgac	ctgaatggaa	gccggcggca	cctcgctaac	ggattcacca	9240
ctccaagaa	ıt tggagccaat	caattcttgc	ggagaactgt	gaatgcgcaa	accaaccctt	9300
ggcagaaca	t atccatcgcg	tccgccatct	ccagcagccg	cacgcggcgc	atctcgggca	9360
gcgttgggt	c ctggccacgg	gtgcgcatga	tcgtgctcct	gtcgttgagg	acccggctag	9420
	g ttgccttact					9480
	g ctgcaaaacg					9540
gtgtttcgt	a aagtctggaa	acgcggaagt	cagogocotg	caccattatg	ttccggatct	9600
gcatcgcag	g atgctgctgg	ctaccctgtg	gaacacctac	atctgtatta	acgaagcgct	9660
ggcattgac	c ctgagtgatt	tttctctggt	cccgccgcat	ccataccqcc	agttgtttac	9720
	g ttccagtaac					9780
	t catcggtatc					9840
	c aaacaggaaa					9900
	t ctggagaaac					9960
	c gaccacgctg					10020
	c ctctgacaca					10080
	c agacaagccc					10140
	c cagtcacgta					10200
	g tactgagagt					10260
	c gcatcaggcg					10320
	c ggcgagcggt					10380
	a acgcaggaaa					10440
aaaaaggcc	g cgttgctggc	gtttttccat	aggeteegee	cccctgacga	qcatcacaaa	10500
	t caagtcagag					10560
	a gctccctcgt					10620
	c tcccttcggg					10680
	t aggtcgttcg					10740
	g ccttatccgg					10800
	g cagcagccac					10860
acagagttc	t tgaagtggtg	gcctaactac	ggctacacta	gaaggacagt	atttggtatc	10920
tacactcta	c tgaagccagt	taccttcqqa	aaaagagttg	gtagetettg	atcccccaaa	10980
caaaccacc	g ctggtagcgg	tggtttttt	gtttgcaage	agcagattac	gcgcagaaaa	11040
aaaggatct	c aagaagatcc	tttgatcttt	tctacaaaat	ctgacgctca	gtggaacgaa	11100
aactcacqt	t aagggatttt	ggtcatgaga	ttatcaaaaa	ggatcttcac	ctagatectt	11160
	a tacgactcac					11184
-	_					

<210> 15 <211> 21 <212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 15 agtatcgtgg tagagagctg c	2:
<210> 16	
<211> 36	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer	
<400> 16	
taatacgact cactataggg atgtggctgg agatgc	36

# (19) World Intellectual Property Organization

International Bureau





(43) International Publication Date 1 July 2004 (01.07.2004)

**PCT** 

# (10) International Publication Number WO 2004/055216 A3

(51) International Patent Classification<sup>7</sup>:

C12Q 1/70

(21) International Application Number:

PCT/US2003/039722

(22) International Filing Date:

12 December 2003 (12.12.2003)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/433,303

13 December 2002 (13.12.2002) US

(71) Applicant (for all designated States except US): FOX CHASE CANCER CENTER [US/US]; 333 Cottman Avenue, Philadelphia, PA 19111 (US).

(72) Inventors; and

- (75) Inventors/Applicants (for US only): ZHU, Qing [CN/US]; 1504 Liberty Court, North Wales, PA 19454 (US). GUO, Ju-Tao [CN/US]; 26 Township Line Road, Apt. A7, Elkins Park, PA 19027 (US). SEEGER, Christoph [US/US]; 407 Waring Road, Elkins Park, PA 19027 (US).
- (74) Agents: RIGAUT, Kathleen, D. et al.; Dann, Dorfman, Herrell & Skillman, Suite 2400, 1601 Market Street, Philadelphia, PA 19103-2307 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- (88) Date of publication of the international search report: 5 August 2004

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/39722

IPC(7)	SSIFICATION OF SUBJECT MATTER : C12Q 1/70			,					
According to	US CL: 235/5 According to International Patent Classification (IPC) or to both national classification and IPC								
	DS SEARCHED								
	ocumentation searched (classification system follower 35/5, 235; 435/354, 358, 363	by classific	ation symbols)						
Documentati	Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched								
Electronić d PubMed	ata base consulted during the international search (na	me of data b	ase and, where practicable, so	earch terms used)					
C. DOC	UMENTS CONSIDERED TO BE RELEVANT								
Category *	Citation of document, with indication, where a			Relevant to claim No.					
X	US 5,968,775 A (HOUGHTON et al.) 19 October	1999 (19.10	.1999), col. 2, lines 63-67;	1, 3, 7, 9-13, 15					
A	col. 3, lines 1-5 and 16-28; col. 13, lines 59-67; an	la coi. 14, 11	mes 1-23	2, 4-6, 8, 14					
X	MACEJAK. D.G. Enhanced antiviral effect in cell ribozymes targeting HCV RNA Journ. of Viral He and 402	culture of ty patitis June 2	ype 1 interferon and 2001, Vol. 8, pages 401	16-18					
<u> </u>									
	documents are listed in the continuation of Box C.		See patent family annex.						
≠ S	pecial categories of cited documents:		later document published after the inter date and not in conflict with the applica	tion but cited to understand the					
	defining the general state of the art which is not considered to be lar relevance	•	principle or theory underlying the inven						
"E" earlier ap	plication or patent published on or after the international filing date		document of particular relevance; the c considered novel or cannot be considere when the document is taken alone						
"L" document establish i specified)	which may throw doubts on priority claim(s) or which is cited to the publication date of another citation or other special reason (as	ω <b>Υ</b> " (	document of particular relevance; the c considered to involve an inventive step combined with one or more other such	when the document is					
"O" document									
	"P" document published prior to the international filing date but later than the "&" document member of the same patent family priority date claimed								
	Date of the actual completion of the international search  O3 May 2004 (03.05.2004)  Date of mailing of the international search report  O3 May 2004 (03.05.2004)								
03 May 2004 (03.05.2004)									
Name and mailing address of the ISA/US  Authorized officer  Mail Step BCT. Attra ISA/US									
O3 May 2004 (03.05.2004)  Name and mailing address of the ISA/US  Mail Stop PCT, Atm: ISA/US  Commissioner for Patents  P.O. Box 1450  Authorized officer  James Housel									
	. Box 1450 xandria, Virginia 22313-1450	Telephone	No. (703) 308-0196	1.00					
	Facsimile No. (703) 305-3230								

Form PCT/ISA/210 (second sheet) (July 1998)